MULTIVARIATE ARCH MODELS: FINITE SAMPLE PROPERTIES OF ML ESTIMATORS AND AN APPLICATION TO A LM-TYPE TEST*

Emma M. Iglesias and Garry D.A. Phillips**

WP-AD 2004-09

Corresponding author: Emma M. Iglesias, University of Alicante, Dpt. Fundamentos del Análisis Económico, Campus de San Vicente, 03080 Alicante, Spain. Fax: +34 96 590 38 98 / E-mail: emiglesias@merlin.fae.ua.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
Primera Edición Febrero 2004.
Depósito Legal: V-1048-2004

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

---

* Both authors thank H. Wong for providing us with the Gauss program to simulate the Wong and Li (1997) model. We acknowledge gratefully as well the financial support from an ESRC grant (Award number: T026 27 1238).

** E.M. Iglesias: University of Alicante. G.D.A. Phillips: Cardiff University.

MULTIVARIATE ARCH MODELS: FINITE SAMPLE PROPERTIES OF ML ESTIMATORS AND AN APPLICATION TO A LM-TYPE TEST

Emma M. Iglesias and Garry D.A. Phillips

ABSTRACT

At the present time, there exists an important and growing econometric literature that deals with the application of multivariate-ARCH models to a variety of economic and financial data. However, the properties of the estimation procedures that are used have not yet been fully explored. This paper provides two main new results: the first concerns the large biases and variances that can arise when the ML estimation method is employed in a simple bivariate structure under the assumption of conditional heteroscedasticity; and the second examines how to use these analytical theoretical results to improve the size and the power of a test for multivariate ARCH effects. We analyse two models: one proposed in Wong and Li (1997) (where the disturbances are dependent but uncorrelated) and another proposed by Engle and Kroner (1995) and Liu and Polasek (1999, 2000) (where conditional correlation is allowed through a diagonal representation). We prove theoretically that a relatively large difference between the intercepts in the two conditional variance equations produces, in the first model, very large variances in some of the ML estimators and, in the second, very severe biases in some of the ML estimators of the parameters. Later we use our bias expressions to propose an LM type test of multivariate ARCH effects, showing that the size and the power of the test improve when we allow for bias correction in the estimators, and that the best recommendation in practical applications is always to use the expected hessian version of the LM. We address as well some constraints that should be included in the estimation of the models but which have so far been ignored. Finally, we present a SUR (seemingly unrelated) specification in both models, that provides an alternative way to retrieve the information matrix. We also extend Lumsdaine (1995) results in multivariate framework.

Keywords: Multivariate GARCH, Bias evaluation.

JEL classification: C13, C32.

1 Introduction

The multivariate-ARCH (autoregressive conditional heteroscedastic) model was first introduced by Kraft and Engle (1983), and Bollerslev, Engle and Wooldridge (1988). Since then, new combinations of this specification in the variance equation with different structures in the mean equation have been proposed: for example Baba, Engle, Kraft and Kroner (1991), Harmon (1988), and Engle and Kroner (1995) introduced the theoretical framework of simultaneous equation models and Calzolari and Fiorentini (1994) have considered some cases of non-linear simultaneous equations, while Polasek and Kozumi (1996) proposed the VAR-GARCH structure.

The multivariate model implies that the conditional variance - covariance matrix (H_t) of the disturbances (ε_t) depends on the information set (I_{t-1}). If we assume normality in the conditional distribution (following Engle and Kroner (1995)), the multivariate-GARCH (Generalised-ARCH) model can be written as:

$$\varepsilon_t / I_{t-1} \sim N\left(0, H_t\right)$$

The main problem to be faced in this specification is the relatively large number of parameters that are involved. There are, however, many possible parameterisations for H_t in order to reduce the number of parameters to estimate. We begin by considering the "vech" (vec-half) representation, which in the simple bivariate case becomes:

$$\begin{aligned} vechH_t = \begin{pmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} c_{01} \\ c_{02} \\ c_{03} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{1,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon^2_{2,t-1} \end{pmatrix} \\ + \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{pmatrix} \end{aligned}$$

For the estimation of this model, it is necessary to restrict the number of parameters still further. Another possible specification is the diagonal representation, where each element of the covariance matrix $h_{jk,t}$ is a function, only, of past values of itself and past values of $\varepsilon_{j,t}\varepsilon_{k,t}$. In the case of the bivariate model, it becomes:

$$vechH_t = \begin{pmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} c_{01} \\ c_{02} \\ c_{03} \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{pmatrix} + \begin{pmatrix} g_{11} & 0 & 0 \\ 0 & g_{22} & 0 \\ 0 & 0 & g_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{pmatrix}$$

The drawback is that we must still ensure that H_t is a positive definite matrix for all values of the ε_t, and it can be a difficult task to check this in the previous specifications. This is why Engle and Kroner (1995) proposed a new parameterisation: the BEKK (Baba, Engle, Kraft and Kroner (1991)) representation, where "...it includes all positive definite diagonal representations, and nearly all positive definite vech representations..." (Engle and Kroner (1995)). In the simple bivariate case, it becomes:

$$H_t = C_0^{*\prime} C_0^* + \begin{pmatrix} a_{11}^* & a_{12}^* \\ a_{21}^* & a_{22}^* \end{pmatrix}' \begin{pmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{pmatrix} \begin{pmatrix} a_{11}^* & a_{12}^* \\ a_{21}^* & a_{22}^* \end{pmatrix} + \begin{pmatrix} g_{11}^* & g_{12}^* \\ g_{21}^* & g_{22}^* \end{pmatrix}' H_{t-1} \begin{pmatrix} g_{11}^* & g_{12}^* \\ g_{21}^* & g_{22}^* \end{pmatrix}$$

Hence, this solution is not as restrictive as the diagonal representation, but "...comparing this model to the vech form of the model, we see that this model economises on parameters by imposing restrictions both across and within equations..." (Engle and Kroner (1995)).

Nowadays there exists an extensive literature about multivariate-ARCH models that have been applied to different varieties of data. Most of them use Maximum Likelihood (ML) as the estimation procedure. However, there are relatively few theoretical papers that examine the consequences of doing that.

The relevant part of the (conditional) log-likelihood function in these models is denoted by:

$$L(y,\theta) = \sum_{t=1}^{T} L_t(y_t,\theta) = -\frac{1}{2}\sum_{t=1}^{T} \log|H_t| - \frac{1}{2}\sum_{t=1}^{T} (y_t - \mu_t)' H_t^{-1} (y_t - \mu_t) \qquad (1.1)$$

Liu and Polasek (1999) gave the following representation of the conditional information matrix of the ML estimator ($I(\theta)$) in a general multivariate heteroscedastic model:

$$I(\theta) = \frac{1}{2}\sum_{t=1}^{T} \left(\frac{\partial vechH_t}{\partial\theta'}\right)' D' \left(H_t^{-1} \bigotimes H_t^{-1}\right) D \frac{\partial vechH_t}{\partial\theta'} + \sum_{t=1}^{T} \left(\frac{\partial\mu_t}{\partial\theta'}\right)' H_t^{-1} \frac{\partial\mu_t}{\partial\theta'} \qquad (1.2)$$

(see Liu and Polasek (1999), page 103), where $\mu_t = E(y_t/I_{t-1})$ is an $M \times 1$ conditional mean vector, $H_t = var(y_t/I_{t-1})$ is an $M \times M$ conditional variance matrix, D is the $M^2 \times M(M+1)/2$ duplication matrix and $\bigotimes$ indicates Kronecker product. They argue that this formula corrects the initial work by Wong and Li (1997) who omitted the last expression of the equation. However, Liu and Polasek (1999) themselves introduced an error when they applied this expression to a VAR(1)-VARCH(1) (Vector-AR(1)-Vector-ARCH(1)) model, because they neglected the influence of changes in the parameters in the variance equation on the own disturbance (see Liu and Polasek (1999, page 105)), which illustrates the difficulties of dealing with these models from the theoretical point of view.

Regarding asymptotic theory, Tuncer (1994, 2000), Bauwens and Vandeuren (1995) and Comte and Lieberman (2003) have established the strong consistency of the Quasi-Maximum Likelihood estimator (QMLE) in a simple multivariate-ARCH model. Asymptotic normality is proved provided that the initial state is either stationary or fixed. More recently, Ling and McAleer (2003) have shown the asymptotic normality in a VARMA-GARCH model requiring only the existence of the second-order moment of the unconditional errors, and a finite fourth-order moment of the conditional errors, which represents an important advance. On the other hand, Hafner (2000), analyses the fourth moment in this model. However, these papers do not consider the issue of the sample size needed in order to have confidence in the asymptotic result, and in this paper we provide results which go some way towards addressing this.

In relation to finite samples, in a more recent paper, Liu and Polasek (2000) have compared through Monte Carlo simulation the biases that are generated using the Splus+GARCH program package of MathSoft (1996), the BASEL package of Polasek (1999) and the application of the method of scoring for MLE using the exact information matrix (given previously). The generated biases are seen to be striking, and the Bayesian method seems to be the best alternative (Appendix 1 shows the results they obtain for a VAR(1)-VARCH(1) model). For a sample size of 200 observations, these results show the existence of severe biases, and this is precisely what has motivated the work in our present paper. On the other hand, Wong and Li (1997) reported through Monte Carlo simulation that in their model, the biases in the parameters were very small (see Wong and Li (1997) pages 119-122). In this paper we are interested in a further analysis of these two models.

The plan of the paper is as follows. In the next section we will begin analysing a bivariate model under two specifications that have been proposed in the literature so far: the one given in Wong and Li (1997), where they allow the two disturbances to be dependent but not correlated, and the one proposed in Engle and Kroner (1995) and Liu and Polasek (1999, 2000), where linear dependence between the disturbances is introduced. We provide theoretical results of the $O(T^{-1})$ biases for the ML estimators in each specification under the assumption of conditional heteroscedasticity. We impose the restriction that the variance parameters are zero; hence following an approach that can be found in a number of other studies (see Engle, Hendry and

Trumble (1985) and Linton (1997)). In effect, we consider the case where conditional heteroscedasticity is assumed when, in fact, it is absent. For easy of manipulation, we assume as well the intercept in the mean equation to be known, although, more complicated structures could be analysed following the same methodology. We prove how in the Wong and Li (1997) model the variances for some estimators can be large when there is a relatively large difference between the intercepts in the variance equations (they only showed results for cases when the intercept parameters had very similar numerical values). On the other hand, in the second model, we show that this large difference in the intercepts can produce very large biases in some of the MLEs. We show as well theoretically how in the Wong and Li (1997) and in the Liu and Polasek (1999, 2000) models some assumptions should be imposed for the ML estimator to be well-defined. We provide evidence that the biases can be very different depending on both the structure we impose on the model, and on the combinations of the parameters we study. We also analyse some invariance properties extending the Lumsdaine (1995) work in univariate framework, and we propose a SUR (Seemingly unrelated) specification in both models, that provides an alternative way to find the information matrix under the null hypothesis of no-ARCH effects. Later, in Section 3, we prove how these bias results can be used to improve the size and the power of a test for multivariate ARCH effects. Finally, Section 4 concludes.

2 A theoretical approach of a bivariate-ARCH model

2.1 Case 1: Allowing for dependent but uncorrelated disturbances

We begin by analysing the framework proposed in Wong and Li (1997) for the variance equation, where the model is specified as:

$$y_t = \beta + \varepsilon_t \tag{2.1}$$

where $y_t = (y_{1t}, y_{2t})'$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$, $E(\varepsilon_t) = 0$, and we assume the intercept vector $\beta = (\beta_{10}\beta_{20})'$ to be known. The conditional variance equation follows the structure (we assume conditional normality):

$H_t = \begin{pmatrix} h_{11t} & 0 \\ 0 & h_{22t} \end{pmatrix}$

where:

$$h_{11t} = E\left(\varepsilon_{1t}^2 / I_{t-1}\right) = \alpha_0 + \alpha_1 \varepsilon_{1t-1}^2 + \alpha_2 \varepsilon_{2t-1}^2 \tag{2.2}$$

$$h_{22t} = E\left(\varepsilon_{2t}^2 / I_{t-1}\right) = \gamma_0 + \gamma_1 \varepsilon_{1t-1}^2 + \gamma_2 \varepsilon_{2t-1}^2 \tag{2.3}$$

Expressions (2.2) and (2.3) can be re-written as:

$$\varepsilon_{1t}^2 = \alpha_0 + \alpha_1 \varepsilon_{1t-1}^2 + \alpha_2 \varepsilon_{2t-1}^2 + \eta_{1t}$$

$$\varepsilon_{2t}^2 = \gamma_0 + \gamma_1 \varepsilon_{1t-1}^2 + \gamma_2 \varepsilon_{2t-1}^2 + \eta_{2t}$$

where, due to the uncorrelatedness of the epsilons:

$$E(\eta_{1t}) = E(\eta_{2t}) = 0; \quad E(\eta_{1t}\eta_{2t}) = 0$$

$$E\left(\eta_{1t}^2\right) = E\left(2h_{11t}^2\right); \; E\left(\eta_{2t}^2\right) = E\left(2h_{22t}^2\right)$$

After some algebra, we find:

$$E\left(\varepsilon_{1t}^2\right) = \frac{\alpha_0(1-\gamma_2) + \alpha_2\gamma_0}{(1-\gamma_2)(1-\alpha_1) - \gamma_1\alpha_2}; \qquad E\left(\varepsilon_{2t}^2\right) = \frac{\gamma_0(1-\alpha_1) + \gamma_1\alpha_0}{(1-\gamma_2)(1-\alpha_1) - \gamma_1\alpha_2}$$

From the above we may deduce the following restrictions on the variance equation parameters:

$$\gamma_2 < 1; \; \alpha_1 < 1; \; (1-\gamma_2)(1-\alpha_1) - \gamma_1\alpha_2 > 0$$

Our objective is to analyse the ML biases of $O(T^{-1})$ in this simple model under the assumption that we specify the conditional variance structure given in (2.2) and (2.3) when, in fact, in the true model we have $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$. This type of assumption, together with the fact that β is considered as known, has been imposed in many of the theoretical analyses that have been carried out in univariate ARCH models so far (eg. Engle, Hendry and Trumble (1985), Linton (1997) and Iglesias and Phillips (2003)), and it facilitates, especially here, the analysis and the interpretation of the results.

There are many papers that show, in univariate ARCH and GARCH models, expressions for the conditional and unconditional moments of the disturbance, but in multivariate-ARCH models we have not found any such previous work. So, we will begin by finding, in our simple setting, the moment expressions that relate the structure of ε_{1t} and ε_{2t}, simply by using the relationships given above:

LEMMA 2.1: *Under the specification of a heteroscedastic disturbance vector $\varepsilon_t = (\varepsilon_{1t}\varepsilon_{2t})'$ that follows (2.1), (2.2) and (2.3), we can establish the following results when $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$:*

1. $E(\varepsilon_{1t}^4) = 3\alpha_0^2$
2. $E(\varepsilon_{2t}^4) = 3\gamma_0^2$
3. $E(\varepsilon_{1t}^6) = 15\alpha_0^3$

4. $E\left(\varepsilon_{2t}^{6}\right) = 15\gamma_0^3$

5. $E\left(\varepsilon_{1t}^{2}\varepsilon_{2t}^{2}\right) = \alpha_0\gamma_0$

6. $E\left(\varepsilon_{1t}^{4}\varepsilon_{2t}^{2}\right) = 3\gamma_0\alpha_0^2$

7. $E\left(\varepsilon_{1t}^{2}\varepsilon_{2t}^{4}\right) = 3\alpha_0\gamma_0^2$

The methodology we will use has been proposed by Cox and Snell (1968), where they showed that for independent, but not necessarily identically distributed observations, the bias (b) of the MLE of β $\left(\widehat{\beta}\right)$ reduces to:

$$b_s = E\left(\widehat{\beta}_s - \beta_s\right) = \sum_{i,j,l=1}^{p} k^{si}k^{jl}\left\{\frac{1}{2}k_{ijl} + k_{ij,l}\right\} + O\left(T^{-2}\right) \tag{2.4}$$

for $s = 1, ..., p$, where $k_{ij} = E\left(\frac{\partial^2 L}{\partial\beta_i\partial\beta_j}\right)$, $k_{ijl} = E\left(\frac{\partial^3 L}{\partial\beta_i\partial\beta_j\partial\beta_l}\right)$, $k_{ij,l} = E\left(\left(\frac{\partial^2 L}{\partial\beta_i\partial\beta_j}\right)\frac{\partial L}{\beta_l}\right)$, for $i, j, l = 1, ..., p$ (L denotes the Log-likelihood function). The total Fisher Information matrix and its inverse are defined by $K = \{-k_{ij}\}$ and $K^{-1} = \{-k^{ij}\}$ respectively. The formula is valid, even for non-independent observations, provided that all $k's$ are of $O(T)$ (see Cordeiro and McCullagh (1991)), and this justifies the application of the methodology in our case.

In order to proceed to obtain the expectations of the second and third order derivatives, we can follow the matrix differential calculus techniques of Magnus and Neudecker (1991). Liu and Polasek (1999) provided the expression of the information matrix $(I(\theta))$ of a general $VAR(k) - VARCH(q)$ model for $y_t = (y_{1t}, y_{2t}, ..., y_{Mt})$, by specialising (1.2):

$$I(\theta) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}$$

with

$$I_{11} = \sum_{t=1}^{T} W_t' H_t^{-1} W_t, \quad I_{21} = I_{12}' = 0$$

and

$$I_{22} = \frac{1}{2}\sum_{t=1}^{T} V_t' D' \left(H_t^{-1} \bigotimes H_t^{-1}\right) D V_t$$

where

$W_t = (I_M, X_{t-1}, ..., X_{t-k})$, $V_t = (I_N, Z_{t-1}, ..., Z_{t-q})$,
$X_{t-i} = diag\left(y_{1t-i}, y_{2t-i}, ..., y_{Mt-i}\right)'$, for $i = 1, ..., k$,
$Z_{t-j} = diag\left(\varepsilon_{1t-j}^2, \varepsilon_{1t-j}\varepsilon_{2t-j}, ...\varepsilon_{Mt-j}^2\right)$, for $j = 1, ..., q$.

Note that I_M and I_N are an $M \times M$ and an $N \times N$ identity matrices respectively and D is the duplication matrix defined in (1.2).

However, if we follow their analysis (see Liu and Polasek (1999, page 105)), the previous expression is found to be in error because it neglects the dependence of the disturbances in the variance equation with respect to the parameters in the mean equation in the maximisation procedure. Their formula is valid only in the situation where there are no parameters to estimate in the mean equation, which is precisely our case. We extend the work by Liu and Polasek (1999) to include all the cumulants we need for our analysis and Appendix 2 provides the expressions for the second and third order derivatives of (1.1) in our model on applying the differential matrix calculus.

Once we know the expressions of all the k components, we are in the position to apply expression (2.4), obtaining the bias results and the variances (given by the information matrix) presented in the next Theorem:

THEOREM 2.1: *If $y_t = \varepsilon_t$ where $y_t = (y_{1t}, y_{2t})'$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a vector of random variables that has the structure given in (2.2) and (2.3), with $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$, then the biases and the variances of the ML estimators to order T^{-1} are given by:*

$E(\widehat{\alpha}_0 - \alpha_0) = \frac{\alpha_0}{T} + o(T^{-1})$ $\quad E(\widehat{\gamma}_0 - \gamma_0) = \frac{\gamma_0}{T} + o(T^{-1})$

$E(\widehat{\alpha}_1 - \alpha_1) = -\frac{1}{T} + o(T^{-1})$ $\quad E(\widehat{\alpha}_2 - \alpha_2) = o(T^{-1})$

$E(\widehat{\gamma}_1 - \gamma_1) = o(T^{-1})$ $\quad E(\widehat{\gamma}_2 - \gamma_2) = -\frac{1}{T} + o(T^{-1})$

$var(\widehat{\alpha}_0) = \frac{4\alpha_0^2}{T} + o(T^{-1})$ $\quad var(\widehat{\gamma}_0) = \frac{4\gamma_0^2}{T} + o(T^{-1})$

$var(\widehat{\alpha}_1) = \frac{1}{T} + o(T^{-1})$ $\quad var(\widehat{\alpha}_2) = \frac{\alpha_0^2}{T\gamma_0^2} + o(T^{-1})$

$var(\widehat{\gamma}_1) = \frac{\gamma_0^2}{T\alpha_0^2} + o(T^{-1})$ $\quad var(\widehat{\gamma}_2) = \frac{1}{T} + o(T^{-1})$

Proof. Given in Appendix 2. ■

It is interesting to note how, when the intercept parameters α_0 and γ_0 differ substantially, the above model can generate severe and large variances in the ML estimators of the α_2 and γ_1 parameters (at least in one of them). In practical applications that fit a model with this specification to real data, one should be mindful of this fact when interpreting the estimation results.

Table 2.1 shows the standard errors of $O(T^{-1})$, and a comparison with the simulated errors for different combinations of the intercepts of the conditional variance equation, confirming the results shown previously.

Table 2.1: Approximate Standard Errors when we mispecify the multivariate ARCH effects. T= 400.

	$\alpha_0 = 0.81$		$\alpha_0 = 0.04$	
	$\gamma_0 = 0.04$		$\gamma_0 = 0.04$	
α_0	0.081	(0.082)	0.004	(0.004)
α_1	0.050	(0.051)	0.050	(0.050)
α_2	1.012	(1.035)	0.050	(0.051)
γ_0	0.004	(0.004)	0.004	(0.004)
γ_1	0.002	(0.002)	0.050	(0.051)
γ_2	0.050	(0.050)	0.050	(0.050)

Simulated values are given in brackets for 20000 replications

On the other hand, the bias and variances of the ML estimators to $O\left(T^{-1}\right)$ in a univariate ARCH(1) model, $E\left(\varepsilon_t^2/I_{t-1}\right) = \alpha_1 + \alpha_2\varepsilon_{t-1}^2$, when nothing is estimated in the mean equation and under misspecification, are given by (see Engle, Hendry and Trumble (1985) and Iglesias and Phillips (2003)):

$E\left(\widehat{\alpha}_1 - \alpha_1\right) = \frac{\alpha_1}{T} + o\left(T^{-1}\right)$ $\quad$ $E\left(\widehat{\alpha}_2 - \alpha_2\right) = -\frac{1}{T} + o\left(T^{-1}\right)$

$var\left(\widehat{\alpha}_1\right) = \frac{3\alpha_1^2}{T} + o\left(T^{-1}\right)$ $\quad$ $var\left(\widehat{\alpha}_2\right) = \frac{1}{T} + o\left(T^{-1}\right)$

Comparing these biases with those of Theorem 2.1, it is seen that in the new bivariate specification the biases in the parameters that are common have the same structure, while on the other hand, there is a loss of estimation efficiency to $O\left(T^{-1}\right)$ in the intercept parameter estimator, and no gain or loss in efficiency for the estimator of the ARCH parameter.

Extending the work in Lumsdaine (1995), the representation of the relevant part of the log-likelihood involves:

$$L_t = -\frac{1}{2}\left(\log h_{11t} + \log h_{22t} + \frac{\varepsilon_{1t}^2}{h_{11t}} + \frac{\varepsilon_{2t}^2}{h_{22t}}\right)$$

Using the same argument as the one given in Lumsdaine (1995, page 10), we can prove that if α_0 and γ_0 change in the same proportion, the biases and t-statistics in $\widehat{\alpha}_1$, $\widehat{\alpha}_2$, $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ will remain invariant. This result matches with the bias and variance results obtained in Theorem 2.1. However, if α_0 and γ_0 vary in different proportions, the invariance property does not hold.

2.2 Case 2: Allowing for dependent and correlated disturbances

We analyse now the variance-specification proposed by Engle and Kroner (1995) and Liu and Polasek (1999,2000), given by the bivariate model:

$$y_t = \beta + \varepsilon_t \tag{2.5}$$

where $y_t = \left(y_{1t}, y_{2t}\right)'$, $\varepsilon_t = \left(\varepsilon_{1t}, \varepsilon_{2t}\right)'$, $E\left(\varepsilon_t\right) = 0$, and we assume again the intercept vector $\beta = \left(\beta_{10}, \beta_{20}\right)'$ to be known. The variance representation implies a diagonal structure for the disturbances following an ARCH(1) process:

$\varepsilon_t/I_{t-1} \sim N\left(0, H_t\right)$, where $H_t = \begin{pmatrix} h_{11t} & h_{12t} \\ h_{21t} & h_{22t} \end{pmatrix}$ and:

$$\begin{pmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{20} \\ \alpha_{30} \end{pmatrix} + \begin{pmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1}^2 \\ \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}^2 \end{pmatrix} \tag{2.6}$$

Then, it follows that:

$$E\left(\varepsilon_{1t}\varepsilon_{2s}/I_{t-1}\right)=\alpha_{20}+\alpha_{22}\varepsilon_{1t-1}\varepsilon_{2s-1},\ t=s \tag{2.7}$$
$$0 \text{ otherwise}$$

$$E\left(\varepsilon_{1t}^{2}/I_{t-1}\right)=\alpha_{10}+\alpha_{11}\varepsilon_{1t-1}^{2} \tag{2.8}$$

$$E\left(\varepsilon_{2t}^{2}/I_{t-1}\right)=\alpha_{30}+\alpha_{33}\varepsilon_{2t-1}^{2}$$

The unconditional expectations become $E\left(\varepsilon_{1t}^{2}\right)=\frac{\alpha_{10}}{1-\alpha_{11}}$, $E\left(\varepsilon_{2t}^{2}\right)=\frac{\alpha_{30}}{1-\alpha_{33}}$, $E\left(\varepsilon_{1t}\varepsilon_{2t}\right)=\frac{\alpha_{20}}{1-\alpha_{22}}$, and the unconditional correlation coefficient between both disturbances is $\frac{\alpha_{20}\sqrt{(1-\alpha_{11})(1-\alpha_{33})}}{(1-\alpha_{22})\sqrt{\alpha_{10}\alpha_{30}}}$. This implies the restriction that in this model, in order to guarantee that the correlation coefficient is absolutely smaller than 1:

$$\alpha_{20}^{2}\left(1-\alpha_{11}\right)\left(1-\alpha_{33}\right)<\left(1-\alpha_{22}\right)^{2}\left(\alpha_{10}\alpha_{30}\right)$$

When $\alpha_{11}=\alpha_{22}=\alpha_{33}=0$, then $\alpha_{20}^{2}<\alpha_{10}\alpha_{30}$. In addition, $\alpha_{10},\alpha_{30}>0$, while $0<\alpha_{11},\alpha_{33}<1$.

Our objective is again to analyse the biases of $O\left(T^{-1}\right)$ in this simple model when we use the ML estimation procedure, under the assumption that we specify a diagonal structure in the conditional variance, when, in fact, the true model is the one for which we have $\alpha_{11}=\alpha_{22}=\alpha_{33}=0$.

In this case, the moment expressions that relate the structure of ε_{1t} and ε_{2t} are given by:

Lemma 2.2: *Under the specification of a heteroscedastic disturbance vector* $\varepsilon_t=\left(\varepsilon_{1t}\varepsilon_{2t}\right)'$ *that follows (2.5) and (2.6), where* $\alpha_{11}=\alpha_{22}=\alpha_{33}=0$, *we can establish the following results to* $O\left(T^{-1}\right)$:

1. $E\left(\varepsilon_{1t}^{3}\varepsilon_{2t}\right)=3\alpha_{10}\alpha_{20}$
2. $E\left(\varepsilon_{1t}^{2}\varepsilon_{2t}^{2}\right)=2\alpha_{20}^{2}+\alpha_{10}\alpha_{30}$
3. $E\left(\varepsilon_{1t}\varepsilon_{2t}^{3}\right)=3\alpha_{20}\alpha_{30}$
4. $E\left(\varepsilon_{1t}^{3}\varepsilon_{2t}^{3}\right)=9\alpha_{10}\alpha_{20}\alpha_{30}+6\alpha_{20}^{3}$
5. $E\left(\varepsilon_{1t}^{4}\varepsilon_{2t}^{2}\right)=12\alpha_{10}\alpha_{20}^{2}+3\alpha_{10}^{2}\alpha_{30}$
6. $E\left(\varepsilon_{1t}^{2}\varepsilon_{2t}^{4}\right)=3\alpha_{10}\alpha_{30}^{2}+12\alpha_{20}^{2}\alpha_{30}$
7. $E\left(\varepsilon_{1t}^{5}\varepsilon_{2t}\right)=15\alpha_{10}^{2}\alpha_{20}$
8. $E\left(\varepsilon_{1t}\varepsilon_{2t}^{5}\right)=15\alpha_{30}^{2}\alpha_{20}$

Proof. Given in Appendix 3. ■

Using again the methodology proposed by Cox and Snell (1968) (Appendix 4 gives the expressions for the second and third order derivatives), the bias results are given in Theorem 2.2.

THEOREM 2.2: *If $y_t = \varepsilon_t$ where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a vector of random variables that has the structure given in (2.6) with $\alpha_{11} = \alpha_{22} = \alpha_{33} = 0$, then the biases and the variances of the ML estimators to order T^{-1} are given by:*

$$E\left(\widehat{\alpha}_{10} - \alpha_{10}\right) = \frac{\alpha_{10}^2\alpha_{30}\left(\alpha_{10}^2\alpha_{30}^2+\alpha_{10}^2\alpha_{20}^2+\alpha_{20}^2\alpha_{30}^2+2\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{20}^4\right)}{T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)} + o\left(T^{-1}\right)$$

$$var\left(\widehat{\alpha}_{10}\right) = \frac{\alpha_{10}^2\left(3\alpha_{10}^3\alpha_{30}^3+13\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+10\alpha_{10}\alpha_{20}^4\alpha_{30}+2\alpha_{20}^6\right)}{T\left(\alpha_{10}^3\alpha_{30}^3+5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+5\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^6\right)} + o\left(T^{-1}\right)$$

$$E\left(\widehat{\alpha}_{20} - \alpha_{20}\right) = \frac{\alpha_{20}\left(\alpha_{10}^4\alpha_{30}^2+\alpha_{10}^2\alpha_{30}^4+6\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+\alpha_{20}^4\alpha_{10}^2+\alpha_{20}^4\alpha_{30}^2-2\alpha_{20}^6\right)}{2T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)} + o\left(T^{-1}\right)$$

$$var\left(\widehat{\alpha}_{20}\right) = \frac{\left(\alpha_{10}^3\alpha_{30}^3+6\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+5\alpha_{10}\alpha_{20}^4\alpha_{30}+2\alpha_{20}^6\right)}{T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^4\right)} + o\left(T^{-1}\right)$$

$$E\left(\widehat{\alpha}_{30} - \alpha_{30}\right) = \frac{\alpha_{10}\alpha_{30}^2\left(\alpha_{10}^2\alpha_{30}^2+\alpha_{10}^2\alpha_{20}^2+\alpha_{20}^2\alpha_{30}^2+2\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{20}^4\right)}{T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)} + o\left(T^{-1}\right)$$

$$var\left(\widehat{\alpha}_{30}\right) = \frac{\alpha_{10}^2\left(3\alpha_{10}^3\alpha_{30}^3+13\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+10\alpha_{10}\alpha_{20}^4\alpha_{30}+2\alpha_{20}^6\right)}{T\left(\alpha_{10}^3\alpha_{30}^3+5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+5\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^6\right)} + o\left(T^{-1}\right)$$

$$E\left(\widehat{\alpha}_{11} - \alpha_{11}\right) = -\frac{\alpha_{10}\alpha_{30}\left(\alpha_{10}^2\alpha_{30}^2+\alpha_{10}^2\alpha_{20}^2+\alpha_{20}^2\alpha_{30}^2+2\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{20}^4\right)}{T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)} + o\left(T^{-1}\right)$$

$$var\left(\widehat{\alpha}_{11}\right) = \frac{\alpha_{10}^2\alpha_{30}^2\left(\alpha_{10}\alpha_{30}+3\alpha_{20}^2\right)}{T\left(\alpha_{10}^3\alpha_{30}^3+5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+5\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^6\right)} + o\left(T^{-1}\right)$$

$$E\left(\widehat{\alpha}_{22} - \alpha_{22}\right) = -\frac{\left(\alpha_{10}^4\alpha_{30}^2+\alpha_{10}^2\alpha_{30}^4+6\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+\alpha_{10}^2\alpha_{20}^4+\alpha_{30}^2\alpha_{20}^4-2\alpha_{20}^6\right)}{2T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)} + o\left(T^{-1}\right)$$

$$var\left(\widehat{\alpha}_{22}\right) = \frac{\left(\alpha_{10}^2\alpha_{30}^2+\alpha_{20}^4\right)}{T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^4\right)} + o\left(T^{-1}\right)$$

$$E\left(\widehat{\alpha}_{33} - \alpha_{33}\right) = -\frac{\alpha_{10}\alpha_{30}\left(\alpha_{10}^2\alpha_{30}^2+\alpha_{10}^2\alpha_{20}^2+\alpha_{20}^2\alpha_{30}^2+2\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{20}^4\right)}{T\left(\alpha_{10}^2\alpha_{30}^2+4\alpha_{10}\alpha_{20}^2\alpha_{30}+\alpha_{20}^4\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)} + o\left(T^{-1}\right)$$

$$var\left(\widehat{\alpha}_{33}\right) = \frac{\alpha_{10}^2\alpha_{30}^2\left(\alpha_{10}\alpha_{30}+3\alpha_{20}^2\right)}{T\left(\alpha_{10}^3\alpha_{30}^3+5\alpha_{10}^2\alpha_{20}^2\alpha_{30}^2+5\alpha_{10}\alpha_{20}^4\alpha_{30}+\alpha_{20}^6\right)} + o\left(T^{-1}\right)$$

Proof. Given in Appendix 4. ■

In spite of the large and tedious expressions we get, it is important to highlight the utility we can get from them, because they allow us to find approximations to the biases for any combination of parameters, and to discover their evolution. In addition, under the assumption of misspecification of the conditional process, we can use the expressions for bias correction, substituting the true values of the expressions with the estimated ones.

Appendix 5 shows the graphs for the 6 expressions in Theorem 2.2 (we must take account of the fact that α_{20} can be negative, and that we must analyse only the regions where $\alpha_{20}^2 < \alpha_{10}\alpha_{30}$). They show that for some combinations of parameters, the biases can become very important. We confirm the results in Liu and Polasek (2000), in the sense that the biases can be very large in these models -even although our setting is different-, but our findings provide evidence that the biases are only so large for some combinations of parameters. Table 2.2 show how the larger biases are those for the parameters of the ARCH-components, especially when there is a large difference between the intercepts of the two conditional variance equations. For example, the approximate bias of the estimator of α_{22} increases from around -0.004 to -0.197 when the constant terms α_{10} and α_{30} change from being the same and equal at 0.15 to α_{10} being kept constant at 0.15 and increasing α_{30} to 15.

Table 2.2: Biases and variances of $O\left(T^{-1}\right)$ for some different parameter configurations, $\alpha_{10} = 0.15$, $\alpha_{20} = 0.5$ and T=200.

	$\alpha_{30} = 0.15$	$\alpha_{30} = 15$
$E\left(\widehat{\alpha}_{10} - \alpha_{10}\right)$	0.00003	0.00713
$var\left(\widehat{\alpha}_{10}\right)$	0.00022	0.00031
$E\left(\widehat{\alpha}_{20} - \alpha_{20}\right)$	0.00177	0.09837
$var\left(\widehat{\alpha}_{20}\right)$	0.00228	0.01336
$E\left(\widehat{\alpha}_{30} - \alpha_{30}\right)$	0.00003	0.71368
$var\left(\widehat{\alpha}_{30}\right)$	0.00022	3.17670
$E\left(\widehat{\alpha}_{11} - \alpha_{11}\right)$	-0.00022	-0.04757
$var\left(\widehat{\alpha}_{11}\right)$	0.00008	0.00411
$E\left(\widehat{\alpha}_{22} - \alpha_{22}\right)$	-0.00355	-0.19675
$var\left(\widehat{\alpha}_{22}\right)$	0.00368	0.00347
$E\left(\widehat{\alpha}_{33} - \alpha_{33}\right)$	-0.00022	-0.04757
$var\left(\widehat{\alpha}_{33}\right)$	0.00008	0.00411

Once we have found the bias expressions of $O\left(T^{-1}\right)$, we can again extend the work by Lumsdaine (1995) to our model. In this case we need to change α_{10}, α_{20} and α_{30} in the same proportion to get invariance in the bias and t-statistics of $\widehat{\alpha}_{11}$, $\widehat{\alpha}_{22}$ and $\widehat{\alpha}_{33}$. Otherwise, the invariance property becomes invalid (again, this is consistent with the results in Theorem 2).

2.2.1 Special case when the correlation of the disturbances is misspecified

In this case, if we set $\alpha_{20} = 0$, Theorem 2.2 now becomes:

Corollary 2.1: *If $y_t = \varepsilon_t$ where $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is a vector of random variables that has the structure given in (2.6) under misspecification of the conditional correlation ($\alpha_{20} = 0$), then the biases and the variances of the ML estimators to order T^{-1} are given by:*

$$
\begin{array}{ll}
E\left(\widehat{\alpha}_{10}-\alpha_{10}\right)=\frac{\alpha_{10}}{T}+o\left(T^{-1}\right) & E\left(\widehat{\alpha}_{11}-\alpha_{11}\right)=-\frac{1}{T}+o\left(T^{-1}\right) \\
E\left(\widehat{\alpha}_{20}-\alpha_{20}\right)=o\left(T^{-1}\right) & E\left(\widehat{\alpha}_{22}-\alpha_{22}\right)=-\frac{\alpha_{10}^{2}+\alpha_{30}^{2}}{2T\alpha_{10}\alpha_{30}}+o\left(T^{-1}\right) \\
E\left(\widehat{\alpha}_{30}-\alpha_{30}\right)=\frac{\alpha_{30}}{T}+o\left(T^{-1}\right) & E\left(\widehat{\alpha}_{33}-\alpha_{33}\right)=-\frac{1}{T}+o\left(T^{-1}\right) \\
var\left(\widehat{\alpha}_{10}\right)=\frac{3\alpha_{10}^{2}}{T}+o\left(T^{-1}\right) & var\left(\widehat{\alpha}_{11}\right)=\frac{1}{T}+o\left(T^{-1}\right) \\
var\left(\widehat{\alpha}_{20}\right)=\frac{\alpha_{10}\alpha_{30}}{T}+o\left(T^{-1}\right) & var\left(\widehat{\alpha}_{22}\right)=\frac{1}{T}+o\left(T^{-1}\right) \\
var\left(\widehat{\alpha}_{30}\right)=\frac{3\alpha_{30}^{2}}{T}+o\left(T^{-1}\right) & var\left(\widehat{\alpha}_{33}\right)=\frac{1}{T}+o\left(T^{-1}\right)
\end{array}
$$

Proof. In the results given in Theorem 2.2, we set $\alpha_{20}=0$. ■

The expression for the bias of $\widehat{\alpha}_{22}$ is now especially easy to interpret and it is easy too to analyse the effect of a large distance between the two intercepts. On the other hand, the bias and variances of the ML estimators in a univariate ARCH(1) model, when nothing is estimated in the mean equation, were given at the end of Section 2.1. So we see that the effect of imposing a correlation between the disturbances, when in fact it does not exist, again does not affect the bias structure, although on the other hand, this time there is neither gain nor loss in efficiency to the order of the approximation.

2.2.2 A SUR-specification for the multivariate ARCH model

In this section, we find an alternative way of getting the expressions for the information matrix under the null hypothesis of no-ARCH effects. Imposing an ARCH structure in the variance equation, implies a SUR (Seemingly-unrelated Regression) structure for the model. If we would apply this methodology to the previous model (Case 1), the GLS (Generalised Least Squares) estimator coincides with the ML and the OLS (Ordinary Least Squares) estimators so the required approximations could be obtained by examining the equations separately. In case 2, however, analysing the problem in the context of a SUR system of equations is helpful under more general conditions so we illustrate the procedure.

Consider again the model:

$$\varepsilon_{1t}^{2}=\alpha_{10}+\alpha_{11}\varepsilon_{1t-1}^{2}+\eta_{1t}$$

$$\varepsilon_{1t}\varepsilon_{2t}=\alpha_{20}+\alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1}+\zeta_{1t}$$

$$\varepsilon_{2t}^{2}=\alpha_{30}+\alpha_{33}\varepsilon_{2t-1}^{2}+\eta_{2t}$$

where η_{1t} and η_{2t} are innovation processes that are connected through the second equation, providing a variance-covariance matrix of the form:

$$\Psi=E\left\{\left(\begin{array}{c}\eta_{1t}\\ \zeta_{1t}\\ \eta_{2t}\end{array}\right)\left(\begin{array}{ccc}\eta_{1t} & \zeta_{1t} & \eta_{2t}\end{array}\right)\right\}$$

The SUR-system can be written in the form:

$$y = X\beta + u$$

where:

$$y = \begin{pmatrix} \varepsilon_{11}^2 \\ \dots \\ \varepsilon_{1n}^2 \\ \varepsilon_{11}\varepsilon_{21} \\ \dots \\ \varepsilon_{1n}\varepsilon_{2n} \\ \varepsilon_{21}^2 \\ \dots \\ \varepsilon_{2n}^2 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & \varepsilon_{10}^2 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \varepsilon_{1(n-1)}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \varepsilon_{10}\varepsilon_{20} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 1 & \varepsilon_{1(n-1)}\varepsilon_{2(n-1)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \varepsilon_{20}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 1 & \varepsilon_{2(n-1)}^2 \end{pmatrix}$$

$$\beta = \begin{pmatrix} \alpha_{10} \\ \alpha_{11} \\ \alpha_{20} \\ \alpha_{22} \\ \alpha_{30} \\ \alpha_{33} \end{pmatrix}, \quad u = \begin{pmatrix} \eta_{11} \\ \dots \\ \eta_{1n} \\ \zeta_{11} \\ \dots \\ \zeta_{1n} \\ \eta_{21} \\ \dots \\ \eta_{2n} \end{pmatrix}, \quad E\left(\mathrm{u}u'\right) = \Sigma = (\Psi^{-1} \bigotimes I_n)$$

where n is the sample size, and Σ is of order $3n \times 3n$.

In this case the GLS estimator is asymptotically equivalent to the ML estimator so that we can obtain asymptotically (under the null of no-ARCH effects):

$$\beta_{GLS} = \beta_{ML} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}Y$$

and the asymptotic variance-covariance matrix can be retrieved from:

$$\lim_{T\to\infty} var\left(\sqrt{T}\left(\beta_{ML} - \beta\right)\right) = \lim_{T\to\infty} E\left(\frac{1}{T}X'\Sigma^{-1}X\right)^{-1}$$

Hence, the information matrix can be approximated by:

$$-E\left(X'\Sigma^{-1}X\right)$$

This procedure could have been easily extended to multivariate models with higher order ARCH-type disturbances, and even generalised to the multivariate-GARCH.

3 An LM type test allowing for bias correction in the estimators

In this section, we show how the small sample bias approximations for the parameters in the Wong and Li (1997) model (Case 1), can be utilised in a Lagrange Multiplier (LM) test. In particular, since in the LM procedure estimation is conducted only under the null, the bias approximations (which for the variance parameters are found only in the null case) can be employed directly. As was seen in Theorem 2.1 bias approximations were found for the constant terms in the variance equations (2.2) and (2.3); these are nuisance parameters for the LM test on the variance parameters, since they are not subject to the test, and bias corrected estimates for them are easily found. These bias corrected estimates will be employed in the LM test. However, an additional use of the bias approximations for the variance parameters in the null case can also be found. Rather than evaluate these parameters as zero under the null, we may set them at the $O\left(T^{-1}\right)$ biases since the expected values of the ML estimators are not zero but are close to the bias approximation. To analyse the effect of this use of the bias corrections, we shall conduct simulations with the bias corrected constant terms in the LM while setting the parameters under test to zero. Then in further simulations we both use bias corrected estimates for the constant terms and set the parameters under test to their asymptotic bias values. We thus define three cases:

Model 1: The nuisance parameters are replaced with uncorrected ML estimates and the parameters under test are set to zero (M1).

Model 2: The nuisance paramaters are replaced with bias corrected ML estimates with the parameters under test set to zero (M2).

Model 3: The nuisance parameters are replaced with bias corrected ML estimates and the parameters under test are set to their asymptotic bias values (M3).

The LM test takes the form:

$LM = S\left(\theta\right)' V^{-1} S\left(\theta\right)$

where S is the score vector, V is the information matrix and $\theta = \left(\alpha_0, \alpha_1, \alpha_2, \gamma_0, \gamma_1, \gamma_2\right)'$ is the 6×1 vector of unknown parameters.

The null hypothesis that we wish to test is:

$H_0 = \alpha_1, \alpha_2, \gamma_1, \gamma_2 = 0$

There are several variants of the LM test and generally they differ only in the estimator of the information matrix; see for example Amemiya (1985) and Dagenais and Dufour (1991) for some related literature. We may distinguish three types of such estimators; the Outer Product (OP) matrix of the score vector, the Hessian (HES) matrix and the expectation of the Hessian (ExpHES) matrix. A non-operational procedure which we shall examine for comparative purposes, uses the true Hessian (TrueHES) where the actual values of unknown parameters are employed rather than estimates. Each of these four variants of the LM test will be examined in the simulations in the contexts of Models 1-3.

The LM test based upon the expected Hessian is not always available since finding the closed form solution for the expected Hessian may not be possible. In this case, however, it is straightforward. From Wong and Li (1997) we find on using (2.1), (2.2) and (2.3), that we may write:

$S(\theta) = \left(-\frac{1}{2h_{11t}}\left(1-\frac{\varepsilon_{1t}^2}{h_{11t}}\right)dh, -\frac{1}{2h_{22t}}\left(1-\frac{\varepsilon_{2t}^2}{h_{22t}}\right)dh\right)'$

$Hessian(\theta) = \begin{pmatrix} hessian_1 & 0 \\ 0 & hessian_2 \end{pmatrix}$

where:

$hessian_i = -\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{2}\left(\frac{2\varepsilon_{it}^2}{h_{iit}}-1\right)\frac{1}{h_{iit}^2}dhdh'\right], i = 1, 2$

$dh = \left(1, \varepsilon_{1t-1}^2, \varepsilon_{2t-1}^2\right)'$

On taking expectations through Hessian(θ) we have:

$$\text{ExpHES}(\theta)=\begin{pmatrix} -\frac{T}{2\alpha_0^2} & -\frac{T}{2\alpha_0} & -\frac{T\gamma_0}{2\alpha_0^2} & 0 & 0 & 0 \\ -\frac{T}{2\alpha_0} & -\frac{3T}{2} & -\frac{T\gamma_0}{2\alpha_0} & 0 & 0 & 0 \\ -\frac{T\gamma_0}{2\alpha_0^2} & -\frac{T\gamma_0}{2\alpha_0} & -\frac{3T\gamma_0^2}{2\alpha_0^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{T}{2\gamma_0^2} & -\frac{T\alpha_0}{2\gamma_0^2} & -\frac{T}{2\gamma_0} \\ 0 & 0 & 0 & -\frac{T\alpha_0}{2\gamma_0^2} & -\frac{3T\alpha_0^2}{2\gamma_0^2} & -\frac{T\alpha_0}{2\gamma_0} \\ 0 & 0 & 0 & -\frac{T}{2\gamma_0} & -\frac{T\alpha_0}{2\gamma_0} & -\frac{3T}{2} \end{pmatrix}$$

with inverse:

$$(ExpHES(\theta))^{-1} = \begin{pmatrix} -\frac{4\alpha_0^2}{T} & \frac{\alpha_0}{T} & \frac{\alpha_0^2}{T\gamma_0} & 0 & 0 & 0 \\ \frac{\alpha_0}{T} & -\frac{1}{T} & 0 & 0 & 0 & 0 \\ \frac{\alpha_0^2}{T\gamma_0} & 0 & -\frac{\alpha_0^2}{T\gamma_0^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{4\gamma_0^2}{T} & \frac{\gamma_0^2}{T\alpha_0} & \frac{\gamma_0}{T} \\ 0 & 0 & 0 & \frac{\gamma_0^2}{T\alpha_0} & -\frac{\gamma_0^2}{T\alpha_0^2} & 0 \\ 0 & 0 & 0 & \frac{\gamma_0}{T} & 0 & -\frac{1}{T} \end{pmatrix}$$

We thus have four variants of the LM test. Their size and power are examined in a set of 60000 simulation experiments. First the test sizes are examined for sample sizes T=50, 100, 200 and 500 where the nuisance parameters are set to $\alpha_0 = 0.81$ and $\gamma_0 = 0.04$. This choice of parameter values was made to ensure that the small sample biases were not trivial. In the simulations, to examine the power of the tests we considered two sets of values for the variance parameters: (i) $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.16$, and (ii) $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.49$. The first of this represents a moderate departure from the null whereas the second lies close to the stationarity bound and so is a relatively extreme departure.

The results on the test size are given in Table 3.1 and for size-adjusted power in Table 3.2. The first clear result we find is that of the bad size properties in small samples for the HES LM test (see Table 3.1), because it is clearly over-sized, even at T=500, in marked contrast to the other tests. The OP and the ExpHES have better size properties, although when we check the size-adjusted power of the tests (Table 3.2) it reveals the lack of power of the OP test for finite samples. At the more extreme alternative the ExpHES and the TrueHES tests have power close to unity at all sample sizes. From the results, the first recommendation in practical applications, is always to use the ExpHES to test for multivariate ARCH effects. Once we have selected the ExpHES, we can concentrate on the selection among Model 1, Model 2 or Model 3. Model 3 seems to have better size properties than Model 1, and Model 2 has better size properties for samples larger than T=100. Analysing the TrueHES, again the one that presents the best size properties is Model 3. If we consider the size-adjusted power, we observe how the test power in Model 2 and 3 always improves on that of Model 1 (the same occurs when we analyse the TrueHES case). So the overall conclusion from the simulations is that, of the operational tests, only ExpHES performs well. Its size is approximately correct even at T=50 while it has high power against both the moderate and extreme alternatives at all sample sizes considered. It even dominates the non-operational TrueHES test for the moderate alternative and has comparable but slightly less power for the extreme alternative. Hence, our overall recommendation must be to use the ExpHES test while bias correcting all the ML estimates, or at least, the ones that are not restricted under the null.

Table 3.1: Size Results based on 5% critical values

	OP			HES			ExpHES			TrueHES		
	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
T=500	0.038	0.038	0.039	0.086	0.081	0.086	0.058	0.056	0.052	0.055	0.059	0.052
T=200	0.047	0.048	0.051	0.146	0.141	0.152	0.057	0.057	0.048	0.062	0.065	0.052
T=100	0.054	0.048	0.054	0.144	0.134	0.146	0.058	0.061	0.044	0.063	0.072	0.053
T=50	0.043	0.040	0.048	0.101	0.095	0.109	0.055	0.062	0.042	0.066	0.080	0.054

The results are based on 60000 Monte Carlo replications under the null of no-ARCH effects. $\alpha_0 = 0.81$ and $\gamma_0 = 0.04$.

Table 3.2: Power Results based on 5% critical values size-adjusted

	OP			HES			ExpHES			TrueHES		
When the alternative hypothesis is $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.16$												
	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
T=500	0.940	0.942	0.932	0.996	0.996	0.996	1.000	1.000	1.000	1.000	1.000	1.000
T=200	0.244	0.250	0.218	0.029	0.032	0.029	1.000	1.000	1.000	0.961	0.966	0.962
T=100	0.062	0.079	0.063	0.018	0.020	0.020	0.979	0.979	0.979	0.852	0.857	0.853
T=50	0.040	0.047	0.042	0.025	0.027	0.027	0.815	0.820	0.832	0.708	0.708	0.709
When the alternative hypothesis is $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0.49$												
T=500	0.837	0.844	0.846	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
T=200	0.537	0.541	0.526	0.719	0.754	0.689	1.000	1.000	1.000	1.000	1.000	1.000
T=100	0.101	0.143	0.087	0.030	0.075	0.015	0.999	1.000	0.999	0.999	0.999	0.999
T=50	0.013	0.018	0.009	0.009	0.009	0.007	0.966	0.969	0.966	0.981	0.982	0.982

The results are based on 60000 Monte Carlo replications. $\alpha_0 = 0.81$ and $\gamma_0 = 0.04$.

4 Conclusions

In this paper we have provided some theoretical evidence of the severe biases and large variances that result from unconstrained-ML estimation of a simple bivariate-ARCH model under misspecification of the conditional heteroscedasticy processes. When we analyse the model in Wong and Li (1997), we find that some of the estimators can have large variances if the difference between the intercepts in the model is relatively large. In the case of the Engle and Kroner (1995) and Liu and Polasek (1999, 2000) specification, we find that a large difference between the intercepts can produce large biases in the estimators of the ARCH-terms for some combinations of parameters. We also find a restriction among the parameters of this model ($\alpha_{20}^2 (1-\alpha_{11})(1-\alpha_{33}) < (1-\alpha_{22})^2 (\alpha_{10}\alpha_{30})$ using the notation in the paper) that should be considered in the estimation. We show too how a SUR representation allows one to easily find the information matrix under the null of no-ARCH effects and how the work given in Lumsdaine (1995) can be extended to the multivariate case. We believe that the possibility of extreme biases and variances should be taken into account in practical applications when ML is used as the estimation procedure in this model. In the last section of the paper we use our bias results to improve, albeit slightly, the size and power of an LM type test for multivariate ARCH effects by bias-correcting the estimators of the parameters. The general recommendation from the simulation results in this paper, is always to use in practical applications the expected Hessian form of the LM test and bias correct all the ML estimators, or at least, the ones that are not restricted under the null. The extension of the results in this paper to more general structures in the mean and in the conditional variance-covariance matrix are subject of future research.

Appendix 1:

Comparisons*of estimates from S-Plus+GARCH, BASEL and MLE with Squared Distances

Parameters	True Values	Splus+G	Basel	MLE
β_{10}	0.10	0.1223657	0.1086234	0.1397224
β_{20}	0.30	0.2004163	0.2678540	0.2855772
β_{11}	-0.40	-0.5246969	-0.5083313	-0.5427349
β_{22}	0.05	0.0493810	0.0158123	0.0404794
α_{10}	0.15	1.0107307	0.0505836	0.2498541
α_{20}	0.05	-0.0001644	0.0519461	0.1158166
α_{30}	0.10	0.6928885	0.1056807	0.2946794
α_{11}	0.01	-0.1427515	0.0304871	0.1397502
α_{22}	0.07	-0.0180374	0.0378279	0.1243003
α_{33}	0.05	0.1011727	0.0279547	0.0774030

*Table extracted from Liu and Polasek (2000), page 5. The beta parameters correspond to a mean equation of the form:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

While the alpha parameters are given by the same structure that we use in expression (2.6).

Appendix 2

The proof of Theorem 2.1, implies the use of expression (2.4) to find the k_{ij}, the k_{ijl} and the $k_{ij,l}$ components. Using differential matrix calculus, defining $H_t^{-1} = \begin{pmatrix} h^{11t} & h^{12t} \\ h^{21t} & h^{22t} \end{pmatrix}$, and assuming the parameter vector to be $(\alpha_0, \gamma_0, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$, we obtain:

SECOND ORDER DERIVATIVES

	evaluation		evaluation		evaluation		evaluation
k_{11}	$-\frac{T\left(h^{11t}\right)^2}{2}$	k_{12}	0	k_{13}	$-\frac{T\varepsilon_{1t-1}^2\left(h^{11t}\right)^2}{2}$	k_{14}	$-\frac{T\left(h^{11t}\right)^2\varepsilon_{2t-1}^2}{2}$
k_{15}	0	k_{16}	0	k_{22}	$-\frac{T\left(h^{22t}\right)^2}{2}$	k_{23}	0
k_{24}	0	k_{25}	$-\frac{T\left(h^{22t}\right)^2\varepsilon_{1t-1}^2}{2}$	k_{26}	$-\frac{T\left(h^{22t}\right)^2\varepsilon_{2t-1}^2}{2}$	k_{33}	$-\frac{T\left(h^{11t}\right)^2\varepsilon_{1t-1}^4}{2}$
k_{34}	$-\frac{T\left(h^{11t}\right)^2\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2}{2}$	k_{35}	0	k_{36}	0	k_{44}	$-\frac{T\left(h^{11t}\right)^2\varepsilon_{2t-1}^4}{2}$
k_{45}	0	k_{46}	0	k_{55}	$-\frac{T\left(h^{22t}\right)^2\varepsilon_{1t-1}^4}{2}$	k_{56}	$-\frac{T\left(h^{22t}\right)^2\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2}{2}$
k_{66}	$-\frac{T\left(h^{22t}\right)^2\varepsilon_{2t-1}^4}{2}$						

THIRD ORDER DERIVATIVES

	evaluation		evaluation		evaluation
k_{111}	$2T\left(h^{11t}\right)^3$	k_{112}	0	k_{113}	$2T\left(h^{11t}\right)^3\varepsilon_{1t-1}^2$
k_{114}	$2T\left(h^{11t}\right)^3\varepsilon_{2t-1}^2$	k_{115}	0	k_{116}	0
k_{212}	0	k_{213}	0	k_{214}	0
k_{215}	0	k_{216}	0	k_{222}	$2T\left(h^{22t}\right)^3$
k_{223}	0	k_{224}	0	k_{225}	$2T\left(h^{22t}\right)^3\varepsilon_{1t-1}^2$
k_{226}	$2T\left(h^{22t}\right)^3\varepsilon_{2t-1}^2$	k_{133}	$2T\left(h^{11t}\right)^3\varepsilon_{1t-1}^4$	k_{134}	$2T\left(h^{11t}\right)^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2$
k_{135}	0	k_{136}	0	k_{233}	0
k_{234}	0	k_{235}	0	k_{236}	0
k_{144}	$2T\left(h^{11t}\right)^3\varepsilon_{2t-1}^4$	k_{145}	0	k_{146}	0
k_{244}	0	k_{245}	0	k_{246}	0
k_{155}	0	k_{156}	0	k_{166}	0
k_{255}	$2T\left(h^{22t}\right)^3\varepsilon_{1t-1}^4$	k_{256}	$2T\left(h^{22t}\right)^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2$	k_{266}	$2T\left(h^{22t}\right)^3\varepsilon_{2t-1}^4$
k_{333}	$2T\left(h^{11t}\right)^3\varepsilon_{1t-1}^6$	k_{334}	$2T\left(h^{11t}\right)^3\varepsilon_{1t-1}^4\varepsilon_{2t-1}^2$	k_{335}	0
k_{336}	0	k_{434}	$2T\left(h^{11t}\right)^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^4$	k_{435}	0
k_{436}	0	k_{444}	$2T\left(h^{11t}\right)^3\varepsilon_{2t-1}^6$	k_{445}	0
k_{446}	0	k_{335}	0	k_{356}	0
k_{455}	0	k_{456}	0	k_{366}	0
k_{466}	0	k_{555}	$2T\left(h^{22t}\right)^3\varepsilon_{1t-1}^6$	k_{556}	$2T\left(h^{22t}\right)^3\varepsilon_{1t-1}^4\varepsilon_{2t-1}^2$
k_{656}	$2T\left(h^{22t}\right)^3\varepsilon_{1t-1}^2\varepsilon_{2t-1}^4$	k_{666}	$2T\left(h^{22t}\right)^3\varepsilon_{2t-1}^6$		

The Cox and Snell (1968) expressions that are required (apart from the second order derivatives, and the third order derivatives previously given), once we evaluate them when $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$, are:

	eval.		eval.		eval.		eval.		
$\frac{1}{2}k_{111} + k_{11,1}$	0	$\frac{1}{2}k_{112} + k_{11,2}$	0	$\frac{1}{2}k_{113} + k_{11,3}$	0	$\frac{1}{2}k_{114} + k_{11,4}$	0	$\frac{1}{2}k_{115} + k_{11,5}$	0
$\frac{1}{2}k_{116} + k_{11,6}$	0	$\frac{1}{2}k_{211} + k_{21,1}$	0	$\frac{1}{2}k_{212} + k_{21,2}$	0	$\frac{1}{2}k_{213} + k_{21,3}$	0	$\frac{1}{2}k_{214} + k_{21,4}$	0
$\frac{1}{2}k_{215} + k_{21,5}$	0	$\frac{1}{2}k_{216} + k_{21,6}$	0	$\frac{1}{2}k_{221} + k_{22,1}$	0	$\frac{1}{2}k_{222} + k_{22,2}$	0	$\frac{1}{2}k_{223} + k_{22,3}$	0
$\frac{1}{2}k_{224} + k_{22,4}$	0	$\frac{1}{2}k_{225} + k_{22,5}$	0	$\frac{1}{2}k_{226} + k_{22,6}$	0	$\frac{1}{2}k_{131} + k_{13,1}$	$-\frac{T}{2\alpha_0^2}$	$\frac{1}{2}k_{132} + k_{13,2}$	$-\frac{T}{2\alpha_0\gamma_0}$
$\frac{1}{2}k_{133} + k_{13,3}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{134} + k_{13,4}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{135} + k_{13,5}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{136} + k_{13,6}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{231} + k_{23,1}$	0
$\frac{1}{2}k_{232} + k_{23,2}$	0	$\frac{1}{2}k_{233} + k_{23,3}$	0	$\frac{1}{2}k_{234} + k_{23,4}$	0	$\frac{1}{2}k_{235} + k_{23,5}$	0	$\frac{1}{2}k_{236} + k_{23,6}$	0
$\frac{1}{2}k_{141} + k_{14,1}$	$-\frac{T\gamma_0}{2\alpha_0^3}$	$\frac{1}{2}k_{142} + k_{14,2}$	$-\frac{T}{2\alpha_0^2}$	$\frac{1}{2}k_{143} + k_{14,3}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{144} + k_{14,4}$	$-\frac{T\gamma_0^2}{2\alpha_0^3}$	$\frac{1}{2}k_{145} + k_{14,5}$	$-\frac{T}{2\alpha_0}$
$\frac{1}{2}k_{146} + k_{14,6}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{241} + k_{24,1}$	0	$\frac{1}{2}k_{242} + k_{24,2}$	0	$\frac{1}{2}k_{243} + k_{24,3}$	0	$\frac{1}{2}k_{244} + k_{24,4}$	0
$\frac{1}{2}k_{245} + k_{24,5}$	0	$\frac{1}{2}k_{246} + k_{24,6}$	0	$\frac{1}{2}k_{151} + k_{15,1}$	0	$\frac{1}{2}k_{152} + k_{15,2}$	0	$\frac{1}{2}k_{153} + k_{15,3}$	0
$\frac{1}{2}k_{154} + k_{15,4}$	0	$\frac{1}{2}k_{155} + k_{15,5}$	0	$\frac{1}{2}k_{156} + k_{15,6}$	0	$\frac{1}{2}k_{251} + k_{25,1}$	$-\frac{T}{2\gamma_0^2}$	$\frac{1}{2}k_{252} + k_{25,2}$	$-\frac{T\alpha_0}{2\gamma_0^3}$
$\frac{1}{2}k_{253} + k_{25,3}$	$-\frac{T\alpha_0}{2\gamma_0^2}$	$\frac{1}{2}k_{254} + k_{25,4}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{255} + k_{25,5}$	$-\frac{T\alpha_0^2}{2\gamma_0^3}$	$\frac{1}{2}k_{256} + k_{25,6}$	$-\frac{T\alpha_0}{2\gamma_0^2}$	$\frac{1}{2}k_{161} + k_{16,1}$	0
$\frac{1}{2}k_{162} + k_{16,2}$	0	$\frac{1}{2}k_{163} + k_{16,3}$	0	$\frac{1}{2}k_{164} + k_{16,4}$	0	$\frac{1}{2}k_{165} + k_{16,5}$	0	$\frac{1}{2}k_{166} + k_{16,6}$	0
$\frac{1}{2}k_{261} + k_{26,1}$	$-\frac{T}{2\alpha_0\gamma_0}$	$\frac{1}{2}k_{262} + k_{26,2}$	$-\frac{T}{2\gamma_0^2}$	$\frac{1}{2}k_{263} + k_{26,3}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{264} + k_{26,4}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{265} + k_{26,5}$	$-\frac{T\alpha_0}{2\gamma_0^2}$
$\frac{1}{2}k_{266} + k_{26,6}$	$-\frac{T}{2\gamma_0}$	$\frac{1}{2}k_{331} + k_{33,1}$	$-\frac{3T}{\alpha_0}$	$\frac{1}{2}k_{332} + k_{33,2}$	$-\frac{3T}{\gamma_0}$	$\frac{1}{2}k_{333} + k_{33,3}$	$-3T$	$\frac{1}{2}k_{334} + k_{33,4}$	$-\frac{3T\gamma_0}{\alpha_0}$
$\frac{1}{2}k_{335} + k_{33,5}$	$-\frac{3T\alpha_0}{\gamma_0}$	$\frac{1}{2}k_{336} + k_{33,6}$	$-3T$	$\frac{1}{2}k_{431} + k_{43,1}$	$-\frac{T\gamma_0}{2\alpha_0^2}$	$\frac{1}{2}k_{432} + k_{43,2}$	$-\frac{T}{2\alpha_0}$	$\frac{1}{2}k_{433} + k_{43,3}$	$-\frac{T\gamma_0}{2\alpha_0}$
$\frac{1}{2}k_{434} + k_{43,4}$	$-\frac{T\gamma_0^2}{2\alpha_0^2}$	$\frac{1}{2}k_{435} + k_{43,5}$	$-\frac{T}{2}$	$\frac{1}{2}k_{436} + k_{43,6}$	$-\frac{T\gamma_0}{2\alpha_0}$	$\frac{1}{2}k_{441} + k_{44,1}$	$-\frac{3T\gamma_0^2}{\alpha_0^3}$	$\frac{1}{2}k_{442} + k_{44,2}$	$-\frac{3T\gamma_0}{\alpha_0^2}$
$\frac{1}{2}k_{443} + k_{44,3}$	$-\frac{3T\gamma_0^2}{\alpha_0^2}$	$\frac{1}{2}k_{444} + k_{44,4}$	$-\frac{3T\gamma_0^3}{\alpha_0^3}$	$\frac{1}{2}k_{445} + k_{44,5}$	$-\frac{3T\gamma_0}{\alpha_0}$	$\frac{1}{2}k_{446} + k_{44,6}$	$-\frac{3T\gamma_0^2}{\alpha_0^2}$	$\frac{1}{2}k_{351} + k_{35,1}$	0
$\frac{1}{2}k_{352} + k_{35,2}$	0	$\frac{1}{2}k_{353} + k_{35,3}$	0	$\frac{1}{2}k_{354} + k_{35,4}$	0	$\frac{1}{2}k_{355} + k_{35,5}$	0	$\frac{1}{2}k_{356} + k_{35,6}$	0
$\frac{1}{2}k_{451} + k_{45,1}$	0	$\frac{1}{2}k_{452} + k_{45,2}$	0	$\frac{1}{2}k_{453} + k_{45,3}$	0	$\frac{1}{2}k_{454} + k_{45,4}$	0	$\frac{1}{2}k_{455} + k_{45,5}$	0

	eval.		eval.		eval.		eval.		
$\frac{1}{2}k_{456} + k_{45,6}$	0	$\frac{1}{2}k_{361} + k_{36,1}$	0	$\frac{1}{2}k_{362} + k_{36,2}$	0	$\frac{1}{2}k_{363} + k_{36,3}$	0	$\frac{1}{2}k_{364} + k_{36,4}$	0
$\frac{1}{2}k_{365} + k_{36,5}$	0	$\frac{1}{2}k_{366} + k_{36,6}$	0	$\frac{1}{2}k_{461} + k_{46,1}$	0	$\frac{1}{2}k_{462} + k_{46,2}$	0	$\frac{1}{2}k_{463} + k_{46,3}$	0
$\frac{1}{2}k_{464} + k_{46,4}$	0	$\frac{1}{2}k_{465} + k_{46,5}$	0	$\frac{1}{2}k_{466} + k_{46,6}$	0	$\frac{1}{2}k_{551} + k_{55,1}$	$-\frac{3T\alpha_0}{\gamma_0^2}$	$\frac{1}{2}k_{552} + k_{55,2}$	$-\frac{3T\alpha_0^2}{\gamma_0^3}$
$\frac{1}{2}k_{553} + k_{55,3}$	$-\frac{3T\alpha_0^2}{\gamma_0^2}$	$\frac{1}{2}k_{554} + k_{55,4}$	$-\frac{3T\alpha_0}{\gamma_0^2}$	$\frac{1}{2}k_{555} + k_{55,5}$	$-\frac{3T\alpha_0^3}{\gamma_0^3}$	$\frac{1}{2}k_{556} + k_{55,6}$	$-\frac{3T\alpha_0^2}{\gamma_0^2}$	$\frac{1}{2}k_{651} + k_{65,1}$	$-\frac{T}{2\gamma_0}$
$\frac{1}{2}k_{652} + k_{65,2}$	$-\frac{T\alpha_0}{2\gamma_0^2}$	$\frac{1}{2}k_{653} + k_{65,3}$	$-\frac{T\alpha_0}{2\gamma_0}$	$\frac{1}{2}k_{654} + k_{65,4}$	$-\frac{T}{2}$	$\frac{1}{2}k_{655} + k_{65,5}$	$-\frac{T\alpha_0^2}{2\gamma_0^2}$	$\frac{1}{2}k_{656} + k_{65,6}$	$-\frac{T\alpha_0}{2\gamma_0}$
$\frac{1}{2}k_{661} + k_{66,1}$	$-\frac{3T}{\alpha_0}$	$\frac{1}{2}k_{662} + k_{66,2}$	$-\frac{3T}{\gamma_0}$	$\frac{1}{2}k_{663} + k_{66,3}$	$-3T$	$\frac{1}{2}k_{664} + k_{66,4}$	$-\frac{3T\gamma_0}{\alpha_0}$	$\frac{1}{2}k_{665} + k_{66,5}$	$-\frac{3T\alpha_0}{\gamma_0}$
$\frac{1}{2}k_{666} + k_{66,6}$	$-3T$								

Appendix 3

Proof. of Lemma 2.2 ■

Under the conditions established in Lemma 2.2, we can describe a projection of ε_{2t} on ε_{1t}:

$$\varepsilon_{2t} = \phi\varepsilon_{1t} + u_t$$

The slope is given by $\phi = \frac{E(\varepsilon_{1t}\varepsilon_{2t})}{E\left(\varepsilon_{1t}^2\right)} = \frac{\alpha_{20}}{\alpha_{10}}$, and we can find the moments that are required for the disturbance to balance the system we are studying, by substituting for the slope parameter, where necessary, in the following:

$E\left(u_t\right) = 0$,

$E\left(u_t^2\right) = E\left(\varepsilon_{2t}^2\right) - 2\phi E\left(\varepsilon_{1t}\varepsilon_{2t}\right) + \phi^2 E\left(\varepsilon_{1t}^2\right) = \frac{\alpha_{10}\alpha_{30} - \alpha_{20}^2}{\alpha_{10}}$ (this confirms again the restriction $\alpha_{10}\alpha_{30} > \alpha_{20}^2$).

$E\left(u_t^3\right) = 0$, $E\left(u_t^4\right) = \frac{3\alpha_{20}^4 - 6\alpha_{20}^2\alpha_{10}\alpha_{30} + 3\alpha_{10}^2\alpha_{30}^2}{\alpha_{10}^2}$

Hence, using the previous expressions and the conditions given in (2.7) and (2.8):

1. $E\left(\varepsilon_{1t}^3\varepsilon_{2t}\right) = E\left(\varepsilon_{1t}^3\left(\phi\varepsilon_{1t} + u_t\right)\right) = 3\alpha_{10}\alpha_{20}$
2. $E\left(\varepsilon_{1t}^2\varepsilon_{2t}^2\right) = E\left(\varepsilon_{1t}^2\left(\phi\varepsilon_{1t} + u_t\right)^2\right) = 2\alpha_{20}^2 + \alpha_{10}\alpha_{30}$
3. $E\left(\varepsilon_{1t}\varepsilon_{2t}^3\right) = E\left(\varepsilon_{1t}\left(\phi\varepsilon_{1t} + u_t\right)^3\right) = 3\alpha_{20}\alpha_{30}$
4. $E\left(\varepsilon_{1t}^3\varepsilon_{2t}^3\right) = E\left(\varepsilon_{1t}^3\left(\phi\varepsilon_{1t} + u_t\right)^3\right) = 9\alpha_{10}\alpha_{20}\alpha_{30} + 6\alpha_{20}^3$
5. $E\left(\varepsilon_{1t}^4\varepsilon_{2t}^2\right) = E\left(\varepsilon_{1t}^4\left(\phi\varepsilon_{1t} + u_t\right)^2\right) = 12\alpha_{10}\alpha_{20}^2 + 3\alpha_{10}^2\alpha_{30}$
6. $E\left(\varepsilon_{1t}^2\varepsilon_{2t}^4\right) = E\left(\varepsilon_{1t}^2\left(\phi\varepsilon_{1t} + u_t\right)^4\right) = 3\alpha_{10}\alpha_{30}^2 + 12\alpha_{20}^2\alpha_{30}$
7. $E\left(\varepsilon_{1t}^5\varepsilon_{2t}\right) = E\left(\varepsilon_{1t}^5\left(\phi\varepsilon_{1t} + u_t\right)\right) = 15\alpha_{10}^2\alpha_{20}$
8. $E\left(\varepsilon_{1t}\varepsilon_{2t}^5\right) = E\left(\varepsilon_{1t}\left(\phi\varepsilon_{1t} + u_t\right)^5\right) = 15\alpha_{30}^2\alpha_{20}$

Appendix 4

The proof of Theorem 2.2, implies the use of expression (2.4) to find the $k_{ij}, k_{ij,l}$ and the k_{ijl} components. Using differential matrix calculus, defining $H_t^{-1} = \begin{pmatrix} h^{11t} & h^{12t} \\ h^{21t} & h^{22t} \end{pmatrix}$, and ordering the parameters as: $\alpha_{10}, \alpha_{20}, \alpha_{30}, \alpha_{11}, \alpha_{22}, \alpha_{33}$, we obtain:

SECOND ORDER DERIVATIVES

	evaluation		evaluation		evaluation		evaluation
k_{11}	$-\frac{T\left(h^{11t}\right)^2}{2}$	k_{12}	$-Th^{11t}h^{12t}$	k_{13}	$-\frac{T\left(h^{12t}\right)^2}{2}$	k_{14}	$-\frac{T\left(h^{11t}\right)^2\varepsilon_{1t-1}^2}{2}$
k_{15}	$-Th^{11t}h^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1}$	k_{16}	$-\frac{T\left(h^{12t}\right)^2\varepsilon_{2t-1}^2}{2}$	k_{22}	$-T\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{23}	$-Th^{12t}h^{22t}$
k_{24}	$-Th^{11t}h^{12t}\varepsilon_{1t-1}^2$	k_{25}	$-T\varepsilon_{1t-1}\varepsilon_{2t-1}\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{26}	$-Th^{22t}h^{12t}\varepsilon_{1t-1}^2$	k_{33}	$-\frac{T\left(h^{22t}\right)^2}{2}$
k_{34}	$-\frac{T\left(h^{12t}\right)^2\varepsilon_{1t-1}^2}{2}$	k_{35}	$-T\varepsilon_{1t-1}\varepsilon_{2t-1}h^{12t}h^{22t}$	k_{36}	$-\frac{T\left(h^{22t}\right)^2\varepsilon_{2t-1}^2}{2}$	k_{44}	$-\frac{T\left(h^{11t}\right)^2\varepsilon_{1t-1}^4}{2}$
k_{45}	$-T\varepsilon_{1t-1}^3\varepsilon_{2t-1}h^{11t}h^{12t}$	k_{46}	$-\frac{T\left(h^{12}\right)^2\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2}{2}$	k_{55}	$-T\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{56}	$-T\varepsilon_{2t-1}^3\varepsilon_{1t-1}h^{22t}h^{12t}$
k_{66}	$-\frac{T\left(h^{22t}\right)^2\varepsilon_{2t-1}^4}{2}$						

THIRD ORDER DERIVATIVES

	evaluation		evaluation		evaluation
k_{111}	$2T\left(h^{11t}\right)^3$	k_{112}	$4T\left(h^{11t}\right)^2 h^{12t}$	k_{113}	$2T\left(h^{12t}\right)^2 h^{11t}$
k_{114}	$2T\left(h^{11t}\right)^3 \varepsilon_{1t-1}^2$	k_{115}	$4T\left(h^{11t}\right)^2 h^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}$	k_{116}	$2Th^{11t}\left(h^{12t}\right)^2 \varepsilon_{2t-1}^2$
k_{212}	$2Th^{11t}\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{213}	$2Th^{12t}\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{214}	$4T\left(h^{11t}\right)^2 h^{12t}\varepsilon_{1t-1}^2$
k_{215}	$2Th^{11t}\varepsilon_{1t-1}\varepsilon_{2t-1}\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{216}	$2Th^{12t}\varepsilon_{2t-1}^2\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{222}	$4Th^{12t}\left(3h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$
k_{223}	$2Th^{22t}\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{224}	$2Th^{11t}\varepsilon_{1t-1}^2\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{225}	$4Th^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}\left(3h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$
k_{226}	$2Th^{22t}\varepsilon_{2t-1}^2\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{133}	$2T\left(h^{12t}\right)^2 h^{22t}$	k_{134}	$2Th^{11t}\left(h^{12t}\right)^2 \varepsilon_{1t-1}^2$
k_{135}	$2Th^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{136}	$2Th^{22t}\left(h^{12t}\right)^2 \varepsilon_{2t-1}^2$	k_{233}	$4T\left(h^{22t}\right)^2 h^{12t}$
k_{234}	$2Th^{12t}\varepsilon_{1t-1}^2\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{235}	$2Th^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1}\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{236}	$4Th^{12t}\left(h^{22t}\right)^2 \varepsilon_{2t-1}^2$
k_{144}	$2T\left(h^{11t}\right)^3 \varepsilon_{1t-1}^4$	k_{145}	$4T\left(h^{11t}\right)^2 h^{12t}\varepsilon_{1t-1}^3\varepsilon_{2t-1}$	k_{146}	$2T\left(h^{12t}\right)^2 h^{11t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2$
k_{244}	$4T\left(h^{11t}\right)^2 h^{12t}\varepsilon_{1t-1}^4$	k_{245}	$2Th^{11t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{246}	$2Th^{12t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$
k_{155}	$2Th^{11t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{156}	$2Th^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}^3\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{166}	$2T\left(h^{12t}\right)^2 h^{22t}\varepsilon_{2t-1}^4$
k_{255}	$4Th^{12t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2\left(3h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{256}	$2Th^{22t}\varepsilon_{1t-1}\varepsilon_{2t-1}^3\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{266}	$4T\left(h^{22t}\right)^2 h^{12t}\varepsilon_{2t-1}^4$
k_{333}	$2T\left(h^{22t}\right)^3$	k_{334}	$2Th^{22t}\left(h^{12t}\right)^2 \varepsilon_{1t-1}^2$	k_{335}	$4T\left(h^{22t}\right)^2 h^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}$
k_{336}	$2T\left(h^{22t}\right)^3 \varepsilon_{2t-1}^2$	k_{434}	$4T\left(h^{12t}\right)^2 h^{11t}\varepsilon_{1t-1}^4$	k_{435}	$2Th^{12t}\varepsilon_{1t-1}^3\varepsilon_{2t-1}\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$
k_{436}	$2T\left(h^{12t}\right)^2 h^{22t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2$	k_{444}	$2T\left(h^{11t}\right)^3 \varepsilon_{1t-1}^6$	k_{445}	$4T\left(h^{11t}\right)^2 h^{12t}\varepsilon_{1t-1}^5\varepsilon_{2t-1}$
k_{446}	$2T\left(h^{12t}\right)^2 h^{11t}\varepsilon_{1t-1}^4\varepsilon_{2t-1}^2$	k_{335}	$2Th^{22t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^2\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{356}	$4T\left(h^{22t}\right)^2 h^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}^3$
k_{455}	$2Th^{11t}\varepsilon_{1t-1}^4\varepsilon_{2t-1}^2\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$	k_{456}	$2Th^{12t}\varepsilon_{1t-1}^3\varepsilon_{2t-1}^3\left(h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{366}	$2T\left(h^{22t}\right)^3 \varepsilon_{2t-1}^4$
k_{466}	$2T\left(h^{12t}\right)^2 h^{22t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^4$	k_{555}	$4Th^{12t}\varepsilon_{1t-1}^3\varepsilon_{2t-1}^3\left(3h^{11t}h^{22t}+\left(h^{12t}\right)^2\right)$	k_{556}	$2Th^{22t}\varepsilon_{1t-1}^2\varepsilon_{2t-1}^4\left(h^{11t}h^{22t}+3\left(h^{12t}\right)^2\right)$
k_{656}	$4T\left(h^{22t}\right)^2 h^{12t}\varepsilon_{1t-1}\varepsilon_{2t-1}^5$	k_{666}	$2T\left(h^{22t}\right)^3 \varepsilon_{2t-1}^6$		

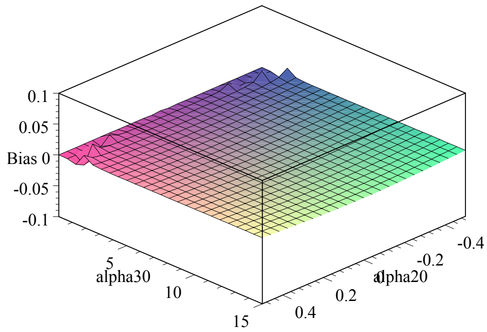
The Cox and Snell (1968) expressions that are required (apart from the second order derivatives, and the third order derivatives previously given), once we evaluate them when $\alpha_{11} = \alpha_{22} = \alpha_{33} = 0$, are:

	evaluation		evaluation
$\frac{1}{2}k_{111} + k_{11,1}$	0	$\frac{1}{2}k_{112} + k_{11,2}$	0
$\frac{1}{2}k_{113} + k_{11,3}$	0	$\frac{1}{2}k_{114} + k_{11,4}$	0
$\frac{1}{2}k_{115} + k_{11,5}$	0	$\frac{1}{2}k_{116} + k_{11,6}$	0
$\frac{1}{2}k_{211} + k_{21,1}$	0	$\frac{1}{2}k_{212} + k_{21,2}$	0
$\frac{1}{2}k_{213} + k_{21,3}$	0	$\frac{1}{2}k_{214} + k_{21,4}$	0
$\frac{1}{2}k_{215} + k_{21,5}$	0	$\frac{1}{2}k_{216} + k_{21,6}$	0
$\frac{1}{2}k_{221} + k_{22,1}$	0	$\frac{1}{2}k_{222} + k_{22,2}$	0
$\frac{1}{2}k_{223} + k_{22,3}$	0	$\frac{1}{2}k_{224} + k_{22,4}$	0
$\frac{1}{2}k_{225} + k_{22,5}$	0	$\frac{1}{2}k_{226} + k_{22,6}$	0
$\frac{1}{2}k_{131} + k_{13,1}$	0	$\frac{1}{2}k_{132} + k_{13,2}$	0
$\frac{1}{2}k_{133} + k_{13,3}$	0	$\frac{1}{2}k_{134} + k_{13,4}$	0
$\frac{1}{2}k_{135} + k_{13,5}$	0	$\frac{1}{2}k_{136} + k_{13,6}$	0
$\frac{1}{2}k_{231} + k_{23,1}$	0	$\frac{1}{2}k_{232} + k_{23,2}$	0
$\frac{1}{2}k_{233} + k_{23,3}$	0	$\frac{1}{2}k_{234} + k_{23,4}$	0
$\frac{1}{2}k_{235} + k_{23,5}$	0	$\frac{1}{2}k_{236} + k_{23,6}$	0
$\frac{1}{2}k_{141} + k_{14,1}$	$\frac{-T\alpha_{10}\alpha_{30}^3}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{142} + k_{14,2}$	$\frac{T\alpha_{20}\alpha_{30}^2(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{143} + k_{14,3}$	$\frac{T\alpha_{20}^2\alpha_{30}^2}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{144} + k_{14,4}$	$\frac{-T\alpha_{10}^2\alpha_{30}^3}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{145} + k_{14,5}$	$\frac{T\alpha_{20}^2\alpha_{30}^2(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{146} + k_{14,6}$	$\frac{T\alpha_{20}^2\alpha_{30}^3}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{241} + k_{24,1}$	$\frac{T\alpha_{10}\alpha_{20}\alpha_{30}^2}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{242} + k_{24,2}$	$\frac{T\alpha_{20}^2\alpha_{30}(\alpha_{30}-\alpha_{10})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{243} + k_{24,3}$	$\frac{-T\alpha_{20}^3\alpha_{30}}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{244} + k_{24,4}$	$\frac{T\alpha_{10}^2\alpha_{20}\alpha_{30}^2}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{245} + k_{24,5}$	$\frac{T\alpha_{20}^3\alpha_{30}(\alpha_{30}-\alpha_{10})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{246} + k_{24,6}$	$\frac{-T\alpha_{20}^3\alpha_{30}^2}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{151} + k_{15,1}$	$\frac{T\alpha_{20}^2\alpha_{30}(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{152} + k_{15,2}$	$\frac{T\alpha_{20}\alpha_{30}\left(\alpha_{30}^2+\alpha_{10}^2-2\alpha_{20}^2\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{153} + k_{15,3}$	$\frac{T\alpha_{20}^2\alpha_{30}(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{154} + k_{15,4}$	$\frac{T\alpha_{10}\alpha_{20}^2\alpha_{30}(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{155} + k_{15,5}$	$\frac{T\alpha_{20}^2\alpha_{30}\left(\alpha_{30}^2+\alpha_{10}^2-2\alpha_{20}^2\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{156} + k_{15,6}$	$\frac{T\alpha_{20}^2\alpha_{30}^2(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{251} + k_{25,1}$	$\frac{T\alpha_{20}(\alpha_{10}-\alpha_{30})\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{252} + k_{25,2}$	$\frac{T\left(2\alpha_{20}^2-\alpha_{10}^2-\alpha_{30}^2\right)\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{253} + k_{25,3}$	$\frac{T\alpha_{20}(\alpha_{30}-\alpha_{10})\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{254} + k_{25,4}$	$\frac{T\alpha_{10}\alpha_{20}(\alpha_{10}-\alpha_{30})\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{255} + k_{25,5}$	$\frac{T\alpha_{20}\left(2\alpha_{20}^2-\alpha_{10}^2-\alpha_{30}^2\right)\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{256} + k_{25,6}$	$\frac{T\alpha_{20}\alpha_{30}(\alpha_{30}-\alpha_{10})\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{161} + k_{16,1}$	$\frac{T\alpha_{20}^4}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{162} + k_{16,2}$	$\frac{T\alpha_{20}^3(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$

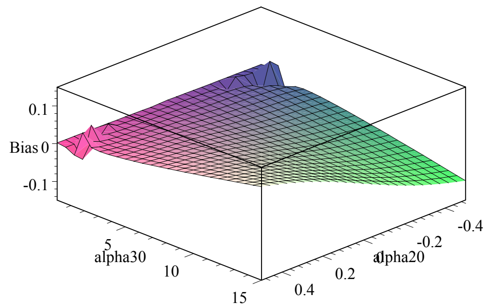
	evaluation		evaluation
$\frac{1}{2}k_{163}+k_{16,3}$	$\frac{-T\alpha_{10}\alpha_{20}^2\alpha_{30}}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{164}+k_{16,4}$	$\frac{T\alpha_{10}\alpha_{20}^4}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{165}+k_{16,5}$	$\frac{T\alpha_{20}^4(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{166}+k_{16,6}$	$\frac{-T\alpha_{10}\alpha_{20}^2\alpha_{30}^2}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{261}+k_{26,1}$	$\frac{-T\alpha_{10}\alpha_{20}^3}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{262}+k_{26,2}$	$\frac{T\alpha_{10}\alpha_{20}^2(\alpha_{10}-\alpha_{30})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{263}+k_{26,3}$	$\frac{T\alpha_{10}^2\alpha_{20}\alpha_{30}}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{264}+k_{26,4}$	$\frac{-T\alpha_{10}^2\alpha_{20}^3}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{265}+k_{26,5}$	$\frac{T\alpha_{10}\alpha_{20}^3(\alpha_{10}-\alpha_{30})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{266}+k_{26,6}$	$\frac{T\alpha_{10}^2\alpha_{20}\alpha_{30}^2}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{331}+k_{33,1}$	0	$\frac{1}{2}k_{332}+k_{33,2}$	0
$\frac{1}{2}k_{333}+k_{33,3}$	0	$\frac{1}{2}k_{334}+k_{33,4}$	0
$\frac{1}{2}k_{335}+k_{33,5}$	0	$\frac{1}{2}k_{336}+k_{33,6}$	0
$\frac{1}{2}k_{431}+k_{43,1}$	$\frac{-T\alpha_{10}\alpha_{20}^2\alpha_{30}}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{432}+k_{43,2}$	$\frac{T\alpha_{20}^3(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{433}+k_{43,3}$	$\frac{T\alpha_{20}^4}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{434}+k_{43,4}$	$\frac{-T\alpha_{10}^2\alpha_{20}^2\alpha_{30}}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{435}+k_{43,5}$	$\frac{T\alpha_{20}^4(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{436}+k_{43,6}$	$\frac{T\alpha_{20}^4\alpha_{30}}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{441}+k_{44,1}$	$\frac{-3T\alpha_{10}^2\alpha_{30}^3}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{442}+k_{44,2}$	$\frac{3T\alpha_{10}\alpha_{20}\alpha_{30}^2(\alpha_{10}-\alpha_{30})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{443}+k_{44,3}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^2}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{444}+k_{44,4}$	$\frac{-3T\alpha_{10}^3\alpha_{30}^3}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{445}+k_{44,5}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^2(\alpha_{10}-\alpha_{30})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{446}+k_{44,6}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^3}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{351}+k_{35,1}$	$\frac{T\alpha_{10}\alpha_{20}^2(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{352}+k_{35,2}$	$\frac{T\alpha_{10}\alpha_{20}\left(\alpha_{30}^2+\alpha_{10}^2-2\alpha_{20}^2\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{353}+k_{35,3}$	$\frac{T\alpha_{10}\alpha_{20}^2(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{354}+k_{35,4}$	$\frac{T\alpha_{10}^2\alpha_{20}^2(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{355}+k_{35,5}$	$\frac{T\alpha_{10}\alpha_{20}^2\left(\alpha_{30}^2+\alpha_{10}^2-2\alpha_{20}^2\right)}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{356}+k_{35,6}$	$\frac{T\alpha_{10}\alpha_{20}^2\alpha_{30}(\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{451}+k_{45,1}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}(3\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{452}+k_{45,2}$	$\frac{3T\alpha_{20}\alpha_{30}\left[\alpha_{10}\left(\left(\alpha_{10}^2-4\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+\alpha_{30}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right)-2\alpha_{20}^4\alpha_{30}\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{453}+k_{45,3}$	$\frac{3T\alpha_{20}^2\alpha_{30}\left[\alpha_{10}^2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)-\alpha_{10}\alpha_{30}\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)+2\alpha_{20}^4\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$	$\frac{1}{2}k_{454}+k_{45,4}$	$\frac{3T\alpha_{10}^2\alpha_{20}^2\alpha_{30}(3\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{455}+k_{45,5}$	$\frac{3T\alpha_{20}^2\alpha_{30}\left[\alpha_{10}\left(\left(\alpha_{10}^2-4\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+\alpha_{30}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right)-2\alpha_{20}^4\alpha_{30}\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$	$\frac{1}{2}k_{456}+k_{45,6}$	$\frac{3T\alpha_{20}^2\alpha_{30}^2\left[\alpha_{10}^2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)-\alpha_{10}\alpha_{30}\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)+2\alpha_{20}^4\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{361}+k_{36,1}$	$\frac{T\alpha_{10}^2\alpha_{20}^2}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{362}+k_{36,2}$	$\frac{T\alpha_{10}^2\alpha_{20}(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{363}+k_{36,3}$	$\frac{-T\alpha_{10}^3\alpha_{30}}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{364}+k_{36,4}$	$\frac{T\alpha_{10}^3\alpha_{20}^2}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{365}+k_{36,5}$	$\frac{T\alpha_{10}^2\alpha_{20}^2(\alpha_{30}-\alpha_{10})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$	$\frac{1}{2}k_{366}+k_{36,6}$	$\frac{-T\alpha_{10}^3\alpha_{30}^2}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$

	evaluation
$\frac{1}{2}k_{461}+k_{46,1}$	$\frac{T\alpha_{20}^2\left[\left(2\alpha_{20}^2\alpha_{30}+\alpha_{10}\alpha_{30}^2+6\alpha_{10}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+6\alpha_{20}^4\alpha_{30}-3\alpha_{10}\alpha_{30}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right]}{4\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{462}+k_{46,2}$	$\frac{T\alpha_{20}^3\left[-\left(2\alpha_{20}^2+\alpha_{10}\alpha_{30}+3\alpha_{10}^2+3\alpha_{30}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+3\alpha_{10}\alpha_{20}^2\alpha_{30}-6\alpha_{20}^4+3\alpha_{10}^2\alpha_{30}^2\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{463}+k_{46,3}$	$\frac{T\alpha_{20}^2\left[\left(2\alpha_{20}^2\alpha_{10}+\alpha_{10}^2\alpha_{30}+6\alpha_{30}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+6\alpha_{20}^4\alpha_{10}-3\alpha_{30}\alpha_{10}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right]}{4\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{464}+k_{46,4}$	$\frac{T\alpha_{10}\alpha_{20}^2\left[\left(2\alpha_{20}^2\alpha_{30}+\alpha_{10}\alpha_{30}^2+6\alpha_{10}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+6\alpha_{20}^4\alpha_{30}-3\alpha_{10}\alpha_{30}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right]}{4\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{465}+k_{46,5}$	$\frac{T\alpha_{20}^4\left[-\left(2\alpha_{20}^2+\alpha_{10}\alpha_{30}+3\alpha_{10}^2+3\alpha_{30}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+3\alpha_{10}\alpha_{20}^2\alpha_{30}-6\alpha_{20}^4+3\alpha_{10}^2\alpha_{30}^2\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{466}+k_{46,6}$	$\frac{T\alpha_{20}^2\alpha_{30}\left[\left(2\alpha_{20}^2\alpha_{10}+\alpha_{10}^2\alpha_{30}+6\alpha_{30}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+6\alpha_{20}^4\alpha_{10}-3\alpha_{30}\alpha_{10}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right]}{4\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{551}+k_{55,1}$	$\frac{T\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)\left[\left(2\alpha_{20}^2\alpha_{30}+\alpha_{30}^2\alpha_{10}+6\alpha_{10}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+3\alpha_{30}\left(2\alpha_{20}^4-\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{10}^2\alpha_{30}^2\right)\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{552}+k_{55,2}$	$\frac{T\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)\left[3\left(\alpha_{10}\alpha_{20}^3\alpha_{30}+\alpha_{10}^2\alpha_{20}\alpha_{30}^2-\alpha_{20}^5\right)-\alpha_{20}\left(2\alpha_{20}^2+\alpha_{10}\alpha_{30}+3\alpha_{10}^2+3\alpha_{30}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)\right]}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{553}+k_{55,3}$	$\frac{T\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)\left[\left(2\alpha_{20}^2\alpha_{10}+\alpha_{10}^2\alpha_{30}+6\alpha_{30}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+3\alpha_{10}\left(2\alpha_{20}^4-\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{10}^2\alpha_{30}^2\right)\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{554}+k_{55,4}$	$\frac{T\alpha_{10}\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)\left[\left(2\alpha_{20}^2\alpha_{30}+\alpha_{30}^2\alpha_{10}+6\alpha_{10}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+3\alpha_{30}\left(2\alpha_{20}^4-\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{10}^2\alpha_{30}^2\right)\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{555}+k_{55,5}$	$\frac{T\alpha_{20}\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)\left[3\left(\alpha_{10}\alpha_{20}^3\alpha_{30}+\alpha_{10}^2\alpha_{20}\alpha_{30}^2-\alpha_{20}^5\right)-\alpha_{20}\left(2\alpha_{20}^2+\alpha_{10}\alpha_{30}+3\alpha_{10}^2+3\alpha_{30}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)\right]}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{556}+k_{55,6}$	$\frac{T\alpha_{30}\left(\alpha_{20}^2+\alpha_{10}\alpha_{30}\right)\left[\left(2\alpha_{20}^2\alpha_{10}+\alpha_{10}^2\alpha_{30}+6\alpha_{30}\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+3\alpha_{10}\left(2\alpha_{20}^4-\alpha_{10}\alpha_{20}^2\alpha_{30}-\alpha_{10}^2\alpha_{30}^2\right)\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{651}+k_{65,1}$	$\frac{3T\alpha_{20}^2\alpha_{10}\left[\alpha_{30}^2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)-\alpha_{10}\alpha_{30}\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)+2\alpha_{20}^4\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{652}+k_{65,2}$	$\frac{3T\alpha_{20}\alpha_{10}\left[\alpha_{30}\left(\left(\alpha_{30}^2-4\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+\alpha_{10}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right)-2\alpha_{20}^4\alpha_{10}\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{654}+k_{65,4}$	$\frac{3T\alpha_{20}^2\alpha_{10}^2\left[\alpha_{30}^2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)-\alpha_{10}\alpha_{30}\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)+2\alpha_{20}^4\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{655}+k_{65,5}$	$\frac{3T\alpha_{20}^2\alpha_{10}\left[\alpha_{30}\left(\left(\alpha_{30}^2-4\alpha_{20}^2\right)\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)+\alpha_{10}^2\left(\alpha_{10}\alpha_{30}+\alpha_{20}^2\right)\right)-2\alpha_{20}^4\alpha_{10}\right]}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^4}$
$\frac{1}{2}k_{656}+k_{65,6}$	$\frac{3T\alpha_{10}\alpha_{20}^2\alpha_{30}^2(3\alpha_{10}-\alpha_{30})}{2\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{661}+k_{66,1}$	$\frac{3T\alpha_{10}^2\alpha_{20}^2\alpha_{30}}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{662}+k_{66,2}$	$\frac{3T\alpha_{10}^2\alpha_{20}\alpha_{30}(\alpha_{30}-\alpha_{10})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{663}+k_{66,3}$	$\frac{-3T\alpha_{10}^3\alpha_{30}^2}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{664}+k_{66,4}$	$\frac{3T\alpha_{10}^3\alpha_{20}^2\alpha_{30}}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{665}+k_{66,5}$	$\frac{3T\alpha_{10}^2\alpha_{20}^2\alpha_{30}(\alpha_{30}-\alpha_{10})}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$
$\frac{1}{2}k_{666}+k_{66,6}$	$\frac{-3T\alpha_{10}^3\alpha_{30}^3}{\left(\alpha_{10}\alpha_{30}-\alpha_{20}^2\right)^3}$

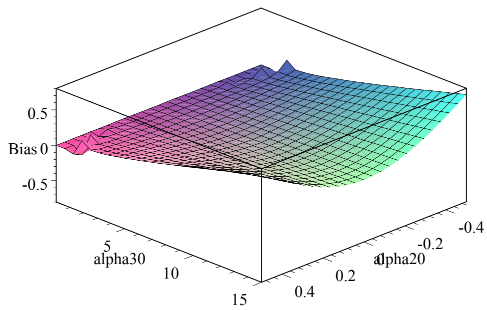
Appendix 5: Graphs of Biases to $O\left(T^{-1}\right)$ in the diagonal case



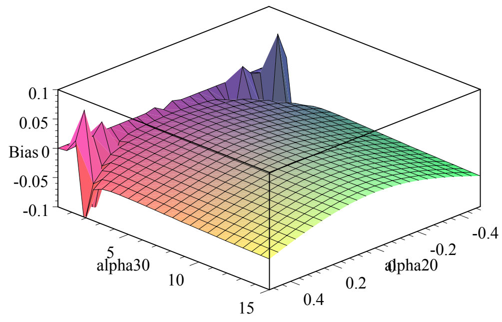
Bias of α_{10} when $\alpha_{10} = 0.15$ for different values of α_{20} and α_{30}. T=200.



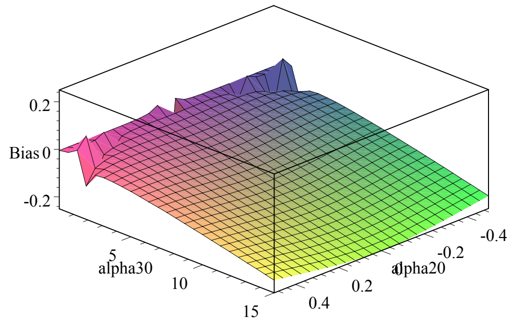
Bias of α_{20} when $\alpha_{10} = 0.15$ for different values of α_{20} and α_{30}. T=200.



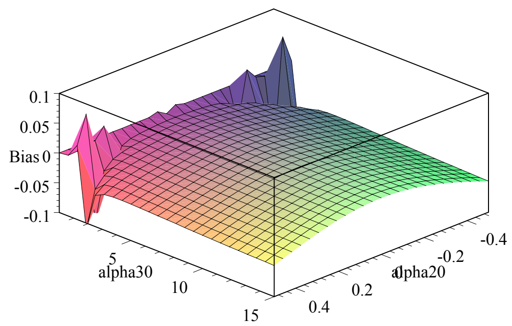
Bias of α_{30} when $\alpha_{10} = 0.15$ for different values of α_{20} and α_{30}. T=200.



Bias of α_{11} when $\alpha_{10} = 0.15$ for different values of α_{20} and α_{30}. T=200.



Bias of α_{22} when $\alpha_{10} = 0.15$ for different values of α_{20} and α_{30}. T=200.



Bias of α_{33} when $\alpha_{10} = 0.15$ for different values of α_{20} and α_{30}. T=200.

References

[1] Amemiya, T. (1985), *Advanced Econometrics*. Harvard University Press, Cambridge.

[2] Baba, Y., R. F. Engle, D. F. Kraft and K. F. Kroner (1991), Multivariate Simultaneous Generalised ARCH, University of California, San Diego: Department of Economics, Discussion Paper No. 89-57.

[3] Bauwens, L. and J.-P. Vandeuren (1995), On the Weak Consistency of the Quasi-Maximum Likelihood Estimator in VAR Models with BEKK-GARCH(1,1) Errors, CORE Discussion Paper 9538.

[4] Bollerslev, T., R. F. Engle and J. M. Wooldridge (1988), A Capital Asset Pricing Model with Time-Varying Covariances, *Journal of Political Economy* 96, 1, 116-131.

[5] Calzolari, G. and G. Fiorentini (1994), Conditional Heteroscedasticity in Nonlinear Simultaneous Equations, EUI Working Paper ECO NO. 94/44.

[6] Comte F. and O. Lieberman (2003), Asymptotic Theory for Multivariate GARCH Processes, *Journal of Multivariate Analysis* 84, 1, 61-84.

[7] Cordeiro, G. M. and P. McCullagh (1991), Bias Correction in Generalised Linear Models, *Journal of the Royal Statistical Society Series B*, 248-275.

[8] Cox, D. R. and E. J. Snell (1968), A General Definition of Residuals, *Journal of the Royal Statistical Society Series B* 30, 248-275.

[9] Dagenais, M. and J. M. Dufour (1991), Invariance, Nonlinear Models, and Asymptotic Tests, *Econometrica* 50, 1601-1615.

[10] Engle, R. F., D. F. Hendry and D. Trumble (1985), Small-sample Properties of ARCH Estimators and Tests, *Canadian Journal of Economics* 18, 66-93.

[11] Engle, R. F. and K. F. Kroner (1995), Multivariate Simultaneous Generalised ARCH, *Econometric Theory* 11, 122-150.

[12] Hafner, C. M. (2000), Fourth Moment of Multivariate GARCH Processes, DP80, SFB 373, Humboldt Universität zu Berlin.

[13] Harmon, R. (1988), The simultaneous Equations Model with Generalised Autoregressive Conditional Heteroscedasticity: the SEM-GARCH Model, Washington D. C.: Board of Governors of the Federal Reserve System, International Finance Discussion Papers, No. 322.

[14] Iglesias, E. M. and G. D. A. Phillips (2003), Small Sample Estimation Bias in GARCH Models with Any Number of Exogenous Variables in the Mean Equation, Mimeo.

[15] Kraft, D. F. and R. F. Engle (1983), Autoregressive Conditional Heteroscedasticity in Multiple Time Series, University of California, San Diego: Department of Economics, Unpublished Manuscript.

[16] Ling S. and M. McAleer (2003), Asymptotic Theory for a Vector ARMA-GARCH Model", *Econometric Theory* forthcoming.

[17] Linton, O. (1997), An Asymptotic Expansion in the GARCH(1,1) Model, *Econometric Theory* 13, 558-581.

[18] Liu, S. and W. Polasek (1999), Maximum Likelihood Estimation for the VAR-VARCH Model: A New Approach. In Modelling and Decisions in Economics, Essays in Honour of Franz Ferschl, Leopold-Wildburger, U., Feichtinger, G. and Kistner, K.-P., eds, Physica-Verlag, Heidelberg, pp. 99-113.

[19] Liu, S. and W. Polasek (2000), On Comparing Estimation Methods for VAR-ARCH Models. Working Paper, University of Basel.

[20] Lumsdaine R. L. (1995), Finite-Sample Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models: a Monte Carlo Investigation, *Journal of Business and Economic Statistics* 13, 1, 1-10.

[21] Magnus, J. R. and H. Neudecker (1991), Matrix Differential Calculus with Applications in Statistics and Econometrics, revised edition. John Wiley, Chichester.

[22] MathSoft (1996), S+GARCH User's Manual, Version 1.0. Data Analysis Products Division, Mathsoft, Seattle.

[23] Polasek, W. et al (1999), The BASEL Package, ISO-WWZ, University of Basel.

[24] Polasek, W. and H. Kozumi (1996), The VAR-VARCH Model: A Bayesian Approach, In Lee J. C., Johnson W. O., Zellner A. Modelling and Prediction Honoring Seymour Geisser, Springer New York, 402-422.

[25] Tuncer, R. (1994), Convergence in Probability of the Maximum Likelihood Estimator of a Multivariate ARMA Model with GARCH(1,1) Errors. CREST Working Paper 9441.

[26] Tuncer, R. (2000), Asymptotic Normality of the Maximum Likelihood Estimators of a Multivariate Random Walk with Drift Model having GARCH(1,1) Errors. Paper presented at the ERC-Metu International Conference in Economics IV, Ankara, September 13-16, 2000.

[27] Wong, H. and W. K. Li (1997), On a Multivariate Conditional Heteroscedastic Model, *Biometrika* 84, 1, 111-123.